Supplemental Digital Content* | Methodology | May 2026

Interpretive Diagnostic Error Reduction: Guideline Update

From the College of American Pathologists in Collaboration With the Association of Directors of Anatomic and Subspecialty Pathology

Authors:
Suzanne M. Dintzis, MD, PhD, FCAP
Stephen S. Raab, MD, FCAP
Marisol Hernandez, MLS, MA
Tanja Kalicanin, MLS(ASCP)^{cm}
Christina Lacchetti, MHSc

Dintzis SM, Evans JJ, Hernandez M, Kalicanin T, Lacchetti C, Nakhleh RE, Otis CN, Pantanowitz L, Parkash V, Raab SS. Interpretive Diagnostic Error Reduction: Guideline Update from the College of American Pathologists in Collaboration with the Association of Directors of Anatomic and Subspecialty Pathology. *Arch Pathol Lab Med*. 2026 Apr. doi:10.5858/arpa.2026-0016-CP. Epub 2026 April 28.

*The supplemental digital content was not copyedited by the *Archives of Pathology & Laboratory Medicine*.

College of American Pathologists | 325 Waukegan Rd. | Northfield, IL 60093 | 800-323-4040 | cap.org

GUIDELINE DEVELOPMENT METHODS

Panel Composition

The College of American Pathologists (CAP) in collaboration with the Association of Directors of Anatomic and Subspecialty Pathology (ADASP) convened an expert and advisory panel (EP/AP) consisting of members with experience and expertise in quality, surgical pathology, cytopathology and community pathology. Members included practicing pathologists and a contracted methodologist. The CAP approved the appointment of the project chair(s) and panel members. ADASP provided official representation for the expert panel.

The roles of each panel are described in the Evidence-based Guideline Development Methodology Manual (Methodology Manual).

Conflict of Interest (COI) Policy

Prior to acceptance on the expert or advisory panel, potential members completed the CAP conflict of interest (COI) disclosure process, whose policy and form require disclosure of material financial interest in, or potential for benefit of significant value from, the guideline's development or its recommendations 24 months prior through the time of publication. The potential members completed the COI disclosure form, listing any relationship that could be interpreted as constituting an actual, potential, or apparent conflict. A complete description of the COI policy is available in the online Methodology Manual.

Everyone was required to disclose conflicts prior to beginning and continuously throughout the project's timeline. The expert panel assessed as having no relevant conflicts of interest. EP members' disclosed conflicts are listed in the Appendix of the manuscript.

The CAP provided funding for the administration of the project; no industry funds were used in the development of the guideline. All panel members volunteered their time and were not compensated for their involvement, except for the contracted methodologist.

Project Scope and Outcomes of Interest

The EP approved the following scope to develop evidence-based recommendations to address the overarching question what are the most effective ways to reduce diagnostic errors in Anatomic Pathology?

The EP approved the key questions for the systematic evidence review.

Systematic Evidence Review

The objective of the systematic evidence review (SER) was to identify articles that provided data to inform the recommendations for interpretive diagnostic error reduction. If of sufficient quality, findings from this review would provide an evidence-base to support the recommendations of the guideline. The scope of the SER and the key questions with the PICO elements (Population, Intervention, Comparator, Outcome(s)) were established by the EP in consultation with the methodologist prior to beginning the literature search.

Detailed key questions including the PICO are included in Supplemental Table 1.

Search and Selection

Detailed literature searches were constructed using controlled vocabulary and keywords for concepts derived from the PICO elements defined at the onset of the project based upon the key questions. All search strategies were reviewed by a second medical librarian using the Press Review of Electronic Search Strategies (PRESS) statement for systematic reviews.[1] All search results were deduplicated using reference management software following published methods.[2] The literature search strategies and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram are included as Supplemental Figures 1 and 2.

Selection at all levels was based on the predetermined inclusion/exclusion criteria:

Studies were selected for inclusion in the systematic review of evidence if they met the following criteria: 1) Peer reviewed articles published since January 1st, 2015, and 2) Study designs that included: observational studies, randomized clinical trials (RCTs); comparative studies. Articles were excluded from the systematic review if: 1) they were published prior to 2015; 2) they were editorials, letters, commentaries, narrative reviews, abstracts, animal studies, and opinions; 3) the full text article was not available in English; 4) they did not address at least one key question; 5) they focused primarily on clinical pathology studies; 6) they focused on issues related to competency use of checklists, standardized language, taxonomy, or formatting; 7) they focused on frozen section diagnoses only; 8) they did not contain a second pathology review or 9) they did not address the outcomes of interest.

A detailed project protocol was developed and registered on PROSPERO, an international database of prospectively registered systematic reviews (Registration number CRD42020153268).[3]

Data Extraction & Management

The data elements from an included article/document were extracted by one reviewer into standard data formats and tables developed using the systematic review database software, DistillerSR (DistillerSR Inc., Ottawa, Canada); a second reviewer confirmed accuracy and completeness. Any discrepancies in data extraction were resolved by discussion between the co-chairs and the methodologist. A bibliographic database was established in EndNote (Clarivate Analytics, Philadelphia, PA) to track all literature identified and reviewed during the study.

Assessing Quality and Risk of Bias

An assessment of the quality of the evidence was performed for all retained studies following application of the inclusion and exclusion criteria. Using this method, studies deemed to be of low quality would not be excluded from the systematic review but would be retained, and their methodological strengths and weaknesses discussed where relevant. To define an overall study quality rating for each included study, validated study-type specific tools were used to assess the risk of bias, plus additional important quality features were extracted. Specific details for each study type are outlined below.

- Systematic Reviews (SRs) and Meta-analyses questions were assessed as per the Assessing the Methodological Quality of Systematic Reviews (AMSTAR) tool[4]

- Single-arm non-randomized phase I and II clinical trials (NRCTs), prospective cohort studies (PCS), prospective-retrospective cohort studies (PRCS), retrospective cohort studies (RCS), and case-control studies (CCS) were assessed using the Risk of Bias in Non-randomized Studies of Intervention (ROBINs-I) tool[5]

- Diagnostic studies were assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool [6]

In the following sections, the quantity of the evidence as determined by the number of studies that met our inclusion criteria and were retained, the evidence type as determined by study design, the quality of that evidence as determined by the quality assessment, and its consistency are all reported, both as individual studies and in totality, statement by statement. Definitions of the certainty of evidence are presented in Supplemental Table 2.

A total of 4295 studies comprised the final body of studies included in the SER. Supplemental Figure 1 displays the results of the literature review. All articles were available as discussion or background references. All members of the EP participated in developing draft recommendations,

reviewing open comment feedback, finalizing and approving the final recommendations, and writing/editing of the manuscript.

For an explanation of the Quality assessment methods and the ROB assessment, refer to the Methodology Manual.

Evidence-to-Decision Framework

This allows for a systematic way to document panel members' judgement for each of the recommendations.[7]

Open Comment Period and Organizational Review

An public, open access comment period was held from October 23, 2024 through November 20, 2024 on the CAP Web site www.cap.org for any interested stakeholder to provide feedback on the draft statements. Six draft statements, two demographic questions, and three questions to assess feasibility were posted for peer review. An announcement was sent to the following societies deemed to have interest.

Medical societies:

- Agency for Healthcare Research and Quality (AHRQ)
- American College of Medical Genetics and Genomics (ACMG)
- American Society for Clinical Oncology (ASCO)
- American Society of Hematology (ASH)
- American Society of Clinical Pathology (ASCP)
- American Society of Cytopathology (ASC)
- American Society for Investigative Pathology (ASIP)
- Association for Molecular Pathology (AMP)
- Association of Community Cancer Centers (ACCC)
- Association of Directors of Anatomic and Subspecialty Pathology (ADASP)
- Association of Pathology Chairs (APC)
- Canadian Association of Pathologists (CAP-APC)
- European Society for Medical Oncology (ESMO)
- Papanicolaou Society of Cytology (PSC)
- Quality Initiative in Interpretive Pathology (QIIP) Canadian Partnership Against Cancer
- Society to Improve Diagnoses in Medicine (SIDM)
- United States & Canadian Academy of Pathology (USCAP)
- Digital Pathology Association
- Association for Pathology Informatics
- State Pathology Societies
- European Society of Pathology
- International Academy of Pathology
- Royal College of Pathologists
- Society for Pediatric Pathology
- American Association of Neuropathologists

Patient advocacy groups:

- Institute for Healthcare Improvement (NPSF)
- Partnership Against Cancer

Government and other stakeholders:

- US Food and Drug Administration (FDA)
- Centers for Medicare & Medicaid Services (CMS)
- Centers for Disease Control and Prevention (CDC)
- Veteran's Affairs (VA) and Department of Defense (DOD)
- European Medical Agency/ EMEA

- CFDA

“Agree” and “Disagree” responses were captured for every proposed recommendation. The EP reviewed all the comments. Resolution of all changes was obtained by majority consensus of the panel using nominal group technique (discussions on teleconference webinars, email discussion and multiple edited recommendations) amongst the panel members. The final recommendations were approved by the EP with a formal vote. Neither formal cost analysis nor cost effectiveness models were performed.

Organizational review was instituted to review and approve the guideline. An independent review panel (IRP) representing the Council on Scientific Affairs was assembled to review and approve the guideline for the CAP. The IRP was masked to the expert panel and vetted through the COI process.

Dissemination Plans

The CAP hosts a resource page which includes a link to the manuscript and supplement; a summary of the recommendations, a teaching PowerPoint (Microsoft Corporation, Redmond, WA), and a frequently asked question (FAQ) document, along with other additional tools such as webinar recordings as applicable. The guideline is promoted and presented at various society meetings and distributed to the societies listed in the peer review.

Recommendation Statements

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

Statement 1 is supported by a total of 52 studies, comprising of one meta-analysis,[8] seven prospective cohort studies,[9-15] 40 retrospective observational studies,[16-55] and four cross-sectional diagnostic accuracy studies.[56-59] Refer to Supplemental Table 3-5 for the quality assessment (QA) results and Supplemental Table 6 for the certainty of evidence for studies informing Statement 1.

Establishing routine procedures for case review in anatomic pathology offers clear benefits but also presents potential challenges. Regular review helps uncover interpretive discrepancies and diagnostic errors that might otherwise go unnoticed, improving consistency, supporting better clinical decisions, and ultimately enhancing patient care. It also promotes shared learning, reinforces adherence to evolving diagnostic standards, and fosters a culture of accountability and continuous improvement. However, review processes can increase workload, potentially delay case reporting, and, if not handled thoughtfully, may cause discomfort or tension among colleagues. There may also be added costs, including time diverted from diagnostic duties, administrative burden, or investment in technology and infrastructure. To be effective and sustainable, review systems should be implemented with clear expectations, practical workflows, and a focus on learning and collaboration rather than blame.

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

Statement 2 is supported by a total of 35 studies that investigated how case reviews influence error detection. The evidence base includes two prospective cohort studies,[9, 60] 28 retrospective observational studies,[61, 62, 16, 63, 18-22, 64-68, 28, 29, 69, 70, 42, 43, 48-55] five cross-sectional diagnostic accuracy studies.[71-74, 58] Refer to Supplemental Table 4-5 for the quality assessment (QA) results and Supplemental Table 7 for the certainty of evidence for studies informing Statement 2.

Timely case review allows pathologists to identify and address potential diagnostic discrepancies while there is still an opportunity to influence patient care. When integrated effectively, such reviews can improve diagnostic accuracy, enhance communication with clinical teams, and

reduce the risk of delays or inappropriate treatment. However, time constraints may introduce additional pressure, potentially leading to incomplete or superficial review, especially if resources are limited. Secondary reviews may also create tension or reluctance among colleagues if not approached in a collegial and constructive manner. Establishing clear expectations and fostering a supportive, collaborative environment are essential to ensuring that timely case review contributes meaningfully to diagnostic quality and patient care.

Supplemental Table 1. Key Questions (KQs) and PICO Element

<table>
<tr><td colspan="3">KQ1. Does targeted review (done at either the analytic or the post-analytic phase) of surgical pathology or cytology cases (slides and/or reports) reduce the error frequency (often measured as amended reports) or increase the rate of error detection compared with no review, random review, or usual review procedures?</td></tr>
<tr><td colspan="3">Population:</td></tr>
<tr><td colspan="3">Surgical pathology or cytology cases</td></tr>
<tr><td>Intervention</td><td>Comparator</td><td>Outcomes</td></tr>
<tr><td>• Targeted review</td><td>• No review, random review, usual review</td><td>• Error frequency
• Change in error rate</td></tr>
<tr><td colspan="3">KQ2. What methods of selecting cases for review have been shown to increase/decrease the rate of error detection compared with no review, random review, or usual review procedures?</td></tr>
<tr><td colspan="3">Population:</td></tr>
<tr><td colspan="3">Surgical pathology or cytology cases</td></tr>
<tr><td>Intervention</td><td>Comparator</td><td>Outcomes</td></tr>
<tr><td>• Methods for selecting cases for review</td><td>• No review, random review, usual review</td><td>• Change in error rate</td></tr>
<tr><td colspan="3">KQ2a. Does the use of artificial intelligence increase or decrease the rate of error reduction?</td></tr>
<tr><td colspan="3">Population:</td></tr>
<tr><td colspan="3">Surgical pathology or cytology cases</td></tr>
<tr><td>Intervention</td><td>Comparator</td><td>Outcomes</td></tr>
<tr><td>• Artificial intelligence</td><td>• Standard procedure without Artificial Intelligence</td><td>• Change in error rate</td></tr>
<tr><td colspan="3">KQ2b. Does the use of ancillary studies in specific clinical situations reduce diagnostic error?</td></tr>
<tr><td colspan="3">Population:</td></tr>
<tr><td colspan="3">Surgical pathology or cytology cases</td></tr>
<tr><td>Intervention</td><td>Comparator</td><td>Outcomes</td></tr>
<tr><td>• Ancillary studies</td><td>• No ancillary studies</td><td>• Error reduction
• Precision and/or accuracy</td></tr>
</table>

Supplemental Table 2: Grades for Certainty of Evidence

Designation	Description
High	There is high confidence that available evidence reflects true effect. Further research is very unlikely to change the confidence in the estimate of effect.
Moderate	There is moderate confidence that available evidence reflects true effect. Further research is likely to have an important impact on the confidence in estimate of effect and may change the estimate.
Low	There is limited confidence in the estimate of effect. The true effect may be substantially different from the estimate of the effect.
Very Low	There is very little confidence in the estimate of effect. The true effect is likely to be substantially different from the estimate of effect. Any estimate of effect is very uncertain.

Data derived from Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) Working Group materials.[75]

Supplemental Table 3. Risk of Bias Assessment of Included Systematic Reviews

Study	AMSTAR Assessment											
	A priori design	Duplicate study selection & data extraction	Comprehensive lit search performed	Grey lit used	List included & excluded studies	Characteristics of included studies provided	Quality assessed & documented	Quality used appropriately for conclusion	Methods to combine used appropriately	Publication bias assessed	COI	**AMSTAR SCORE /11**
Gerhard, R. et al,[8] 2015	√	x	x	x	x	√	x	x	x	x	x	2

Abbreviations: AMSTAR, Assessing the Methodological Quality of Systematic Reviews; COI, conflict of interest

Supplemental Table 4. Risk of Bias Assessment of Included Observational Studies

Study	ROBINS-I Assessment							
	Confounding	Patient Selection	Intervention classification	Deviations from intended interventions	Missing data	Measurement of outcomes	Selection of reported result	Overall Risk of Bias
Abro et al,[61]2021	moderate	low	low	low	low	low	moderate	Moderate
Adhya et al,[9] 2019	moderate	moderate	low	low	moderate	moderate	low	Moderate
Agrawal et al,[62] 2020	moderate	low	moderate	moderate	moderate	low	moderate	Moderate
Agrawal et al,[16] 2021	moderate	low	low	low	moderate	moderate	low	Moderate
Ahmad et al,[63] 2016	moderate	moderate	low	low	moderate	moderate	low	Moderate
Al-Adnani et al,[76] 2015	moderate	moderate	low	moderate	moderate	low	moderate	Moderate
Amin et al,[17] 2019	moderate	moderate	moderate	low	low	moderate	moderate	Moderate
Azara et al,[77] 2016	moderate	moderate	low	low	low	moderate	moderate	Moderate
Bailey et al,[18] 2022	moderate	low	moderate	low	moderate	serious	moderate	Serious
Balasubramanian et al,[19] 2015	moderate	moderate	low	low	moderate	moderate	low	Moderate
Banet et al,[20] 2015	moderate	moderate	low	low	moderate	moderate	low	Moderate

Bellevicine et al, [21] 2017	moderate	moderate	low	low	moderate	low	low	Moderate
Bhoyrul et al,[22] 2019	moderate	moderate	low	low	moderate	serious	low	Moderate
Boennelycke et al,[78] 2021	low	moderate	moderate	low	low	low	low	Moderate
Bulten et al,[79] 2021	low	moderate	moderate	low	low	moderate	low	Moderate
Choi et al,[23] 2021	moderate	low	low	moderate	moderate	low	moderate	Moderate
Cocks et al,[64] 2021	moderate	moderate	moderate	low	low	low	moderate	Moderate
Confortini et al,[24] 2016	moderate	moderate	low	moderate	moderate	low	moderate	Moderate
Crescenzi et al,[80] 2020	moderate	low	low	moderate	low	low	moderate	Moderate
de Morais et al,[81] 2020	serious	moderate	low	moderate	serious	low	moderate	Serious
Dessauvagie et al,[82] 2018	moderate	moderate	low	low	moderate	moderate	low	Moderate
El Sharouni et al,[65] 2021	low	low	low	low	low	moderate	low	Low
Endo et al,[25] 2015	moderate	moderate	moderate	low	low	moderate	moderate	Moderate
Galli et al,[26] 2020	moderate	moderate	low	low	moderate	moderate	moderate	Moderate
Gavrielides et al,[83] 2020	moderate	low	low	low	moderate	moderate	moderate	Moderate
Giunchi et al,[27] 2017	serious	moderate	low	moderate	moderate	low	moderate	Serious
Gordetsky et al,[66] 2018	moderate	low	low	low	low	low	moderate	Moderate
Goyal et al,[10] 2016	moderate	low	low	moderate	low	low	moderate	Moderate
Grevenkamp et al,[84] 2017	moderate	moderate	low	moderate	moderate	low	moderate	Moderate
Groen et al,[11] 2017	moderate	moderate	low	low	moderate	low	moderate	Moderate
Gru et al,[12] 2018	moderate	low	low	low	low	low	moderate	Moderate
Hekler et al,[85] 2019	moderate	low	low	moderate	low	low	moderate	Moderate
Hernandez-Prera et al,[86] 2017	moderate	serious	low	moderate	moderate	low	serious	Serious
Hescot et al,[87] 2021	moderate	low	low	moderate	moderate	serious	moderate	Serious

Hohnen et al,[67] 2021	low	moderate	low	low	low	moderate	low	Moderate
Jimeno et al,[13] 2020	moderate	moderate	low	low	moderate	low	moderate	Moderate
Johnson et al,[68] 2021	moderate	moderate	low	low	moderate	low	moderate	Moderate
Kang et al,[88] 2021	moderate	low	low	low	low	moderate	low	Moderate
Khazai et al,[28] 2015	moderate	low	low	low	moderate	low	moderate	Moderate
Kuijpers et al,[29] 2016	low	low	low	moderate	low	serious	low	Moderate
Kuijpers et al,[60] 2015	moderate	low	low	moderate	moderate	low	moderate	Moderate
Layfield et al,[69] 2017	low	low	low	low	low	moderate	low	Low
Layfield et al,[14] 2017	moderate	low	low	moderate	moderate	low	moderate	Moderate
Layfield et al,[30] 2015	moderate	low	moderate	moderate	moderate	low	moderate	Moderate
Lobo et al,[89] 2021	moderate	moderate	low	moderate	moderate	serious	low	Serious
Lohman et al,[31] 2021	serious	low	low	low	moderate	serious	low	Serious
Lucas et al,[90] 2019	moderate	low	low	low	moderate	moderate	low	Moderate
Magalhaes et al,[15] 2018	moderate	low	low	low	moderate	moderate	low	Moderate
Makela et al,[91] 2018	moderate	low	moderate	low	low	moderate	low	Moderate
Marshall et al,[92] 2018	moderate	moderate	moderate	low	moderate	low	low	Moderate
Mastracci et al,[93] 2016	moderate	moderate	low	moderate	moderate	low	moderate	Moderate
Matheus et al,[94] 2019	moderate	low	low	low	moderate	low	moderate	Moderate
Matsuda et al,[95] 2018	low	moderate	low	low	low	moderate	low	Moderate
Mazariegos et al,[96] 2021	moderate	low	moderate	low	moderate	moderate	low	Moderate
Mullin et al,[32] 2015	moderate	low	low	low	low	serious	low	Moderate
Orlando et al,[33] 2016	moderate	low	low	moderate	moderate	low	moderate	Moderate
Packer et al,[97] 2022	moderate	serious	low	low	moderate	serious	moderate	Serious

Patrawala et al,[34] 2016	serious	serious	low	moderate	moderate	serious	moderate	Serious
Peck et al,[35] 2018	moderate	low	low	moderate	moderate	low	moderate	Moderate
Shubhasini et al,[36] 2017	moderate	moderate	low	moderate	moderate	low	moderate	Moderate
Rakha et al,[98] 2017	moderate	low	moderate	low	low	moderate	low	Moderate
Ronchi et al,[37] 2020	moderate	low	low	moderate	moderate	low	moderate	Moderate
Schollenberg et al,[99] 2015	serious	moderate	serious	moderate	moderate	serious	moderate	Serious
Soofi et al,[38] 2015	moderate	moderate	low	moderate	moderate	low	moderate	Moderate
Stewart et al,[39] 2017	serious	moderate	low	moderate	serious	low	moderate	Serious
Strosberg et al,[100] 2018	low	moderate	low	moderate	moderate	serious	moderate	Serious
Szecsei et al,[101] 2021	moderate	low	moderate	low	moderate	moderate	moderate	Moderate
van der Wel et al,[102] 2021	moderate	serious	low	moderate	moderate	low	moderate	Serious
van Santvoort et al,[103] 2020	moderate	moderate	low	moderate	moderate	moderate	low	Moderate
van Seijen et al,[104] 2021	serious	moderate	moderate	moderate	moderate	moderate	low	Serious
Vats et al,[40] 2022	moderate	moderate	low	low	moderate	moderate	moderate	Moderate
Vennalaganti et al,[105] 2017	moderate	low	low	low	low	moderate	low	Moderate
Villa et al,[106] 2018	moderate	moderate	low	low	moderate	low	moderate	Moderate
Villanacci et al,[107] 2018	low	low	low	low	low	low	low	Low
Yoshida et al,[108] 2022	moderate	moderate	low	low	moderate	moderate	moderate	Moderate
Yue et al,[70] 2020	moderate	moderate	low	moderate	moderate	moderate	moderate	Moderate
Henna et al,[109] 2017	moderate	moderate	low	moderate	moderate	moderate	moderate	Moderate
McHugh et al,[41] 2019	moderate	low	low	low	moderate	moderate	low	Moderate
Middleton et al,[42] 2016	moderate	low	low	moderate	moderate	moderate	low	Moderate
Raju et al,[43] 2019	moderate	moderate	moderate	moderate	moderate	moderate	moderate	Moderate

Solivas et al,[44] 2024	moderate	moderate	low	low	low	moderate	low	Moderate
Boler et al,[45] 2022	moderate	moderate	low	low	low	moderate	low	Moderate
Nagose et al,[46] 2022	moderate	moderate	low	low	low	moderate	low	Moderate
Lezcano et al,[47] 2023	low	low	low	low	low	moderate	low	Low
Richards et al,[48] 2023	moderate	moderate	low	low	low	moderate	low	Moderate
Sripodok et al,[49] 2023	low	low	low	low	low	low	low	Low
Woeste et al,[50] 2022	serious	moderate	low	low	moderate	moderate	moderate	Serious
Rakheja et al,[51] 2022	serious	moderate	low	low	moderate	low	moderate	Moderate
Shinohara et al,[52] 2022	moderate	low	low	low	moderate	low	low	Moderate
Torous et al,[53] 2022	serious	moderate	low	low	serious	moderate	moderate	Serious
Gupta et al,[54] 2022	moderate	moderate	low	low	moderate	low	moderate	Moderate
Wellauer et al,[55] 2022	moderate	low	low	low	low	low	low	Low

Abbreviations: ROBINS-I, Risk of Bias in Non-randomized Studies of Intervention

Supplemental Table 5. Risk of Bias Assessment of Included Diagnostic Studies

Study	**QUADAS**			**Overall Risk of Bias**
	Are there concerns that the included patients do not match the review question?	Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Are there concerns that the target condition as defined by the reference standard does not match the review question?	
Elmore et al,[56] 2017	low	low	low	Low
Elmore et al,[71] 2015	low	low	low	Low
Govind et al,[72] 2020	low	moderate	low	Moderate
Piepkorn et al,[57] 2019	low	low-moderate	low	Low
VandenBussche et al,[73] 2015	low	moderate	low	Moderate
Witt et al,[74] 2016	low	low	low	Low
Van Der Wel et al,[58] 2019	low	low	low	Low
Nethanel et al,[59] 2023	low	low	low	Low

Abbreviations: QUADAS, Quality Assessment of Diagnostic Accuracy Studies

Supplemental Table 6. Certainty of Evidence Assessment for Statement 1

Outcome	Number of Studies	Design	Aggregate Risk of Bias	Inconsistency	Indirectness	Imprecision	Other	Importance	Certainty of Evidence Grade for Outcome	Overall Certainty of Evidence Grade for Statement
Kappa	14	Observational	Serious	Not serious	Not serious	Not serious	None	Critical	Moderate	Moderate
Concordance/ Discordance Discrepancies	34	Observational	Serious	Not serious	Not serious	Not serious	None	Critical	Moderate	
Sensitivity Specificity PPV, NPV Accuracy Other	9	1 MA, 8 Observational	Serious	Serious	Not serious	Not serious	None	Critical	Moderate	

Abbreviations: MA, meta-analysis; PPV, positive predictive value; NPV, negative predictive value.
52 References total:[9, 16-24, 56, 25, 26, 8, 27, 10-13, 28, 29, 14, 30, 31, 15, 32-35, 57, 36-43, 58, 44, 59, 45-55] *Kappa*:[22, 24, 25, 27, 10, 11, 13, 30, 33, 57, 45-47, 49] *Concordance/Discordance/Discrepancies*:[17-24, 26, 12, 28, 29, 14, 31, 15, 32-35, 57, 37-40, 42, 58, 44, 59, 48, 50, 52-55] *Sensitivity/Specificity/PPV/NPV/Accuracy/Other*:[9, 16, 56, 8, 36, 41, 43, 51, 54]

Supplemental Table 7. Certainty of Evidence Assessment for Statement 2

Outcome	Number of Studies	Design	Aggregate Risk of Bias	Inconsistency	Indirectness	Imprecision	Other	Importance	Certainty of Evidence Grade for Outcome	Overall Certainty of Evidence Grade for Statement
Kappa	5	Observational	Serious	Not serious	Not serious	Not serious	None	Critical	Moderate	Moderate
Concordance/ Discordance Discrepancies	25	Observational	Serious	Not serious	Not serious	Not serious	None	Critical	Moderate	
Sensitivity Specificity PPV, NPV Accuracy Other	8	Observational	Serious	Not serious	Not serious	Not serious	None	Critical	Moderate	

Abbreviations: PPV, positive predictive value; NPV, negative predictive value.
35 References total:[61, 9, 62, 16, 63, 18-22, 64, 65, 71, 66, 72, 67, 68, 28, 29, 60, 69, 73, 74, 70, 42, 43, 58, 48-55] *Kappa*:[62, 22, 73, 74, 49] *Concordance/Discordance/Discrepancies*:[63, 18-22, 64, 65, 71, 66-68, 28, 29, 60, 69, 73, 70, 42, 58, 48, 50, 52, 53, 55] *Sensitivity/Specificity/PPV/NPV/Accuracy*:[61, 9, 16, 72, 74, 43, 51, 54]

Supplemental Table 8. Targeted Review

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
BREAST						
Agrawal et al,[16] 2021	Pathologist without designated expertise Retrospective	1147 consecutive breast cases	Internal-single institution Non-gyn cytology Surgical pathology, non-Gyn cytopathology	Malignant cytopathology case found malignant on histopathology With ROSE Sensitivity: 92.9% Specificity: 100% PPV: 100% NPV: 94.9% Accuracy: 96.9% AUC: 0.97 Without ROSE Sensitivity: 82.1% Specificity: 99.4% PPV: 99.2% NPV: 86.9% Accuracy: 91.5% AUC: 0.91 Suspicious and malignant cytopathology case found in situ or malignant on histopathology With ROSE Sensitivity: 98.8% Specificity: 99.1% PPV: 98.8% NPV: 99.1% Accuracy: 99.0% AUC: 0.99 Without ROSE Sensitivity: 99.3% Specificity: 99.4% PPV: 99.3% NPV: 99.4% Accuracy: 99.4% AUC: 0.99%		Incorporating the IAC Yokohama System for Reporting Breast FNAB Cytopathology with ROSE can improve early and accurate diagnosis of breast lesions, prevent missed diagnoses, and provide reliable estimates of ROM in a given patient population.
Khazaiet al,[28] 2015	Expert Retrospective	1970 referral for breast cancer cases	External-multiple institution Unknown/Not indicated		Significant discrepancy: 11.47%	A 11.47% significant error rate for cases referred to the Breast service at authors'

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
			Surgical pathology			institution in 2010 was calculated.
Orlando et al,[33] 2016	Expert Retrospective	187 consecutive malignant breast cancer cases	External-multiple institutions Resection; cases with bx only were excluded Surgical pathology		κ (95% CI) ER 0.612 (0.538-0.686), PgR 0.659 (0.580-0.737) Ki67 0.609 (0.534-0.684) Grading 0.669 (0.569-0.769) HER2 0.546 (0.444-0.649) Local vs central discordance rate: HER2 positive (negative to positive): 0.095 (12/126) HER2negative (positive to negative): 0.317 (13/41) ER/PR (positive to negative): 0.125 (18/145) ER/PR (negative to positive): 0.262 (11/42) After changes in HER2 status and hormonal receptors status (negativity vs positivity and vice versa) 23(12.2%) and 33(17.6%) systemic prescriptions were modified respectively	Although kappas were moderate to substantial, the discordance rate had significant treatment impacts.
Soofi et al,[38] 2015	Expert Retrospective	All core needles breast biopsies of patients referred to institution for further management, N=502	External-single institution Biopsy Surgical pathology		Discordance: 20.72% of all cases reviewed (104/502) Major discordance (potential impact- patient could have had a second lesion which required a definitive therapy or patient could have been loss to follow up): 7.97% (40/502) 15 of the 40 patients with a major discordance had a change in management (3% of all cases reviewed) *NOTE: MAJORITY DID NOT HAVE*	All breast biopsies seen in a referral setting should have a secondary expert review before therapeutic measures are initiated.

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
					FOLLOW-UP (25 WERE LOST TO FU) Minor discordance (no impact on patient care) 12.75% (64/502) Major discordance 5/502 malignant cases 16/502 premalignant cases 10/502 cases with biomarker related 6/502 fibroepthelial lesion type cases Higher rate of discordance in borderline lesions	
Boler et al,[45] 2022	Expert Retrospective	70 consecutive breast cases	Internal-single institution FNA Non-GYN cytopathology		Overall agreement between observers: 70.48% Overall free marginal κ = 0.63 Combining "Suspicious" and "Malignant" categories Overall agreement 80.95% Overall free marginal κ = 0.75 Single category vs multiple category scoring difference major discrepancy 5.70%	A few modifications could be made in the IAC Yokohama categories especially merging category 4 and category 5 that will increase inter-observer agreement.
Woeste et al,[50] 2022	Expert Retrospective	263 consecutive breast cases	External-single institution Excisional biopsy Surgical pathology		Discordant diagnosis: 35% Change in tumor subtype: 10% Change in invasive cancer grade: 45%	Although routine expert breast pathology consultation resulted in few changes in clinical and operative management within this cohort, it ultimately served its purpose as a quality assurance measure ensuring that each patient received the correct treatment. Excisional biopsies and benign lesions without atypia may be associated with less diagnostic

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
						discordance. However, more studies are needed to evaluate risk for pathologic discordance after expert pathology consultation for a case-by-case approach to be accepted as an adequate substitute. Expert pathology review remains essential in the multidisciplinary care of all breast cancer patients.
McHugh et al,[41] 2019	Pathologist without designated expertise Retrospective	219 breast cases	Internal-single institution FNAs Surgical pathology	Sensitivity: 85.0% Specificity: 75.0% PPV: 69.0% NPV: 88.0%		FNAC remains a valuable tool in evaluation of palpable breast lesions in resource-limited healthcare settings.
ENDOCRINE						
Bellevicine et al,[21] 2017	Expert Retrospective	34 consecutive thyroid cases	External-multiple institution FNAs Non-GYN cytopathology		Concordance: 64.71% (22/65) Discordance: 35%	Intra-institutional second opinion practice for “indeterminate” thyroid FNA avoids unnecessary surgeries and maximises the detection of malignant cases diagnosed as FN/SFN.
Gerhard et al,[8] 2015	Unknown/Not indicated Retrospective	7154 consecutive thyroid cases, all diagnoses	External-multiple institution Non-GYN cytopathology	Based on cases with histological follow-up PPV SOD = 55.8 RFNA = 37.7% *P* < .001 Accuracy SOD = 67.4% RFNA = 56.0% *P* = .0034		Second-opinion diagnosis is more accurate than repeat fine-needle aspiration
Sripodok et al,[49]2023	Pathologist without designated expertise	All case with borderline, atypia of undetermined significance/FLUS diagnosis and	Internal-single institution FNA Non-GYN cytopathology		Consensus diagnosis and pathologist interpretation κ range 0.674 to 0.898	There are significant interobserver variations, and changes in cytological diagnosis when performing re-evaluation among different

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
	Retrospective	follow up histology, N=72				pathologists using TBSRTC. A second or third opinion should be routinely sought to establish a consensus diagnosis as an adjunctive method for an initial AUS/FLUS diagnosis. The reclassification lowers medical costs and the rate of unnecessary surgery, particularly in patients with cytologically confirmed benign thyroid nodules. Under the condition that the cytological sample is cellularly adequate, preoperative molecular evaluation is a promising method for aiding in the diagnosis of thyroid nodules that requires additional study.
Galli et al,[26] 2020	Expert Retrospective	72 referred cases of malignant thymus cases	External-single institution Biopsy, resection Surgical pathology		Discrepancies: Different type of thymoma: 26.4% Lung cancer reviewed as thymic cancer: 2.8% Lung cancer reviewed as thymoma: 1.4% Carcinoma NOS reviewed as thymic cancer: 11.1% Lymphoma reviewed as thymoma: 1.4%	This study underlines the importance for TETs to get a second pathological diagnosis by an expert pathologist and supports the need for networks on rare cancers.
ENT/HEAD & NECK						
Adhyaet al,[9] 2019	Unknown/Not indicated Prospective	280 consecutive malignant ENT, oral cavity cases	Internal-single institution Cytopathology Non-GYN cytology, Non-GYN surgical pathology	Sensitivity: 98.2% Specificity: 89.3% PPV: 97.3% NPV: 92.6% Accuracy: 96.4%		TIC aids in the diagnosis of oral squamous cell carcinoma but does not replace incisional biopsy. The cytologic examination of imprints of biopsies also helps determine the adequacy of the specimen for histologic examination.

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
Shubhasini et al,[36] 2017	Expert Not reported	100 oral potentially malignant disorders	Internal-single institution Biopsy Oral lesions	Two systems were used, an 8-tier specific diagnosis and a 2-tier low-grade and high grade. Chi square: 2-tier 0.834, $P = 0.36$		There was poor agreement amongst two expert pathologists on the classification of oral potentially malignant disorders, using two different classification systems. It was a slightly better based in a binary system for better concordance. They determined that biomarkers were needed to improve diagnosis.
Richards et al,[48] 2023	Expert Retrospective	23 consecutive borderline, oral verrucous squamous lesion cases	Internal-single institution Biopsy Surgical pathology		Diagnostic concordance between biopsy and resection (*with respect to a definitive diagnosis on biopsy and agreement on resection*): 17%	The rampant interobserver variability in diagnosis of oral cavity squamous lesions is highlighted and the importance of identifying morphologic clues that can aid in correct diagnosis are emphasizes, thereby helping in adequate clinical management.
Choi et al,[23] 2021	Expert Retrospective	397 referred cases of sinonasal malignancy	Internal-single institution Biopsy, excisional biopsy, resection Surgical pathology		Major discrepancies: 24.4% Minor discrepancies: 61.5%	In histologies prone to misdiagnoses, obtaining a second opinion from experienced head and neck pathologists at a high-volume institution may potentially lead to a change in treatment recommendations that could result in improved survival in patients with sinonasal malignancies.
Mullin et al,[32] 2015	Expert Retrospective	All 566 referral cases to an oral and maxillofacial pathology unit	External-single institution Biopsy Surgical pathology		Total discrepancies: 51.24% Major: 10.25%	Expertise in oral and maxillofacial pathology (dentist+ training in OMF) was important to accurate diagnosis of these lesions. Pathologist who do head and neck pathology have variable training/exposure to OMF.

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
GASTROINTESTINAL						
Aminet al,[17] 2019	Expert Retrospective	323 consecutive GI, appendix, borderline cases	External-multiple institution Resection Surgical pathology		Comparing the histology report from the referring centre with the central review report: Total discrepancies: 18% (57/323) Overall diagnostic discrepancies: 12% (39/323) 22 (7%) patients with local LAMN diagnosis changed in review diagnosis to non-neoplastic reactive mucosal changes associated with either a ruptured diverticulum (21 cases) or resolving acute appendicitis (1 case). HAMN on review in 11 cases, local diagnosis LAMN in 8 and adenocarcinoma in 3. GCC identified on review in 5 cases, initially diagnosed as mucinous adenocarcinoma and reclassified as adenocarcinoma ex-GCC (Tang classification)	Results support central review of histopathology slides by experienced pathologists at tertiary referral centres. This practice can prevent inappropriate management. Findings also suggest that pathologists should consider asking for a second opinion in these unusual and often difficult cases when interpretation is not straightforward. There is also a need to clarify diagnostic criteria. This will depend on future studies in which morphological features are correlated with outcome data and molecular profiling.
Goyal et al,[10] 2016	Expert Prospective	20 mixed pancreas cases; Group A: n=12 Group B: n=8	Internal-Multiple institution Non-GYN Cytopathology Unknown/Not reported		HG dysplasia or worse: 66.67% (8/12) LG or intermediate-grade: 8 cases Group A: chance-adjusted agreement κ = 0.65 Group B: the chance-adjusted agreement κ = 0.03	Cytologic recognition of HG dysplasia or worse as HG atypia in pancreatic mucinous cysts has a good degree of interobserver reproducibility among cytopathologists. In contrast, a problematic area with a lack of agreement appears to be the cytologic recognition of LG and IG dysplasia as LG atypia.

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
Jimeno et al,[13] 2020	Pathologist without designated expertise Prospective	Two rounds of diagnosis were carried out during a period of 18 months on 125 mixed colon cases.	External-multiple institution Biopsy Surgical pathology		Overall interobserver agreement: κ=0.73 Pairwise agreement: κ=0.86 Best intraobserver agreement: κ=0.85 Notable differences were seen between the pathologists with a high-volume and low-volume practice (best overall κ=0.61 and 0.41, respectively)	Examining a large volume of samples is the key factor to increase the consistency in the diagnosis and gradation of IBD-related dysplasia.
		Sub-analysis was performed by eliminating the highly prevalent nondysplastic samples, as well as an analysis after observers' grouping			After eliminating the highly prevalent nondysplastic samples, the interobserver agreement Best overall: κ=0.50 Best paired: κ=0.72	
Van Der Wel et al,[58] 2019	Mixture of experts and non-experts Retrospective	55 cases of Barrett's	External-multiple institution Biopsy Surgical pathology		NDBO Discordance 21.2% Major over interpretation: HGD: 28.9% LGD: 58.4% IND: 77.1% Major diagnostic error in 8.80% of all diagnoses	Data provide evidence-based criteria for diagnostic proficiency in Barrett's histopathology
Layfield et al,[30] 2015	Expert Retrospective	40 sequential brushing specimens of atypical or suspicious for malignancy bile or pancreatic ducts	External-multiple institution Non-GYN cytopathology Surgical pathology		Category atypical k=0.21 Category suspicious for malignancy k=0.18	Analysis of agreement shows that interobserver agreement was only slight to fair. Despite the categories "atypical" and "suspicious for malignancy" having distinct risks of malignancy (62% versus 74%), the reproducibility of these categories is relatively poor. A single intermediate category may improve reproducibility over the scheme proposed by the Papanicolaou Society of

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
						Cytopathology while maintaining an ability to stratify risk of malignancy.
GYNECOLOGY						
Banet et al,[20] 2015	Expert Retrospective	237 consecutive pelvic washing cases	External-multiple institution GYN cytopathology GYN cytology		Major discordance (*counted if the diagnosis changed from negative to positive, or vice versa*): 3.80% (9/237) Minor discordance: 12.24% (29/237)	Only a few cases showed major or minor discordance, and of those that were discordant, most involved ovarian specimens. Major discordance between original and consulting diagnoses would have changed treatment in only 2 cases.
Confortini t et al,[24] 2016	Cytologists, peer review process of all abnormal smears by 10 cytologists Prospective	1086 cervical, ASC-US+ and above cases	Internal-single institution GYN cytopathology GYN cytology		κ = 0.62 (95% CI 0.54-0.69) Downgraded from abnormal to negative: 6% (represents overall performance of all 10 cytologists)	Peer review represents an important method of internal quality control in the evaluation and improvement of inter-observer agreement
Groen et al,[11] 2017	Unknown/Not indicated Prospective	100 mixed cervicovaginal cases	Internal-single institution GYN cytopathology GYN cytology		Digital cytology intermodality agreement among 3 observers who had sufficient training in digital pathology Concordance rates: 81% - 90% κ: 0.76 - 0.86 Agreement among other 2 observers who did not have sufficient training in digital pathology Concordance rates: 56% - 57% κ: 0.41 - 0.44	Panoptiq appears to be feasible for the interpretation of cervicovaginal cytology specimens but requires adequate training in digital pathology.
Layfield et al,[14] 2017	Pathologist without designated expertise Prospective	827 cervical biopsy and curettage specimens undergoing weekly consensus conference review.	Internal-multiple institution Biopsy Surgical pathology		Concordance: First month 2014: 93.9% Last month 2015: 99.0% Total Discrepancies: First month 2014: 31.0%	The consensus conference technique appears to be a useful method to reduce intradepartmental diagnostic discrepancies. Both absolute and chance-corrected

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
					Last month 2015: 8.0% Major Discrepancies: First month 2014: 6.0% Last month 2015: 2.0%	agreement are improved by using consensus conferences.
Magalhaes et al,[15] 2018	Expert Unknown/Not indicated	10868 consecutive pap smears, all negative pap smears selected	Internal-single institution GYN cytopathology GYN cytology		Discrepancies (false negatives): 2004: 0.31% 2013: 1.68%	Ten years after the implementation of the 100% rapid review method, improvements were found in the detection of atypia and precursor lesions and in quality indicators.
Stewart et al,[39] 2017	Expert Retrospective	92 consecutive cases from the tumor registry that had dated terminology or where there was diagnostic uncertainty based on the original report or outcome.	External-multiple institution Resection Surgical Pathology		Discrepancies: Total 68.48% High grade serous: 84.6% Endometrioid: 50.0% LG serous: 30.0% Clear cell 60.0% Mucinous carcinoma 60.0%	This study illustrated specific pitfalls in the diagnosis of ovarian epithelial neoplasia and helped to maintain the accuracy of the Western Australia Cancer Registry.
LUNG						
Endo et al,[25] 2015	Review based on comparison to reference cases with highest interpretive agreement. Retrospective	150 cases of sputum cytology for consensus diagnosis.	External-multiple institution Non-GYN cytology Non-Gyn cytopathology		Kappa: <10 years' experience: 0.27 >10 years' experience: 0.45 Non-respiratory specialty: 0.31 Respiratory specialty: 0.56	Only cytotechnologists with considerable experience in this field make accurate diagnoses on sputum cytology.
SKIN						
Balasubramanian et al,[19] 2015	Expert Retrospective	3006 consecutive skin cases	Internal-single institution Biopsy Surgical pathology		Concordant: 59.81% (1798/3006) Partially concordant: 7.58% (228/3006) Discordant: 30.90% (929/3006)	Several deficits were identified that need to be rectified to improve the diagnostic accuracy of skin biopsy. A portion of discordant slides showing features compatible

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
						with the disease when reviewed by a pathologist and dermatologist together emphasizes the importance of a joint review by both in doubtful cases.
Bhoyrul et al,[22] 2019	Expert Retrospective	341 consecutive melanoma cases	External-multiple institution Biopsy, Excisional biopsy Surgical pathology		Inter-observer variability: 6.74% (23/341) Change management: 2.93% (10/341)	Disagreements in the pathological assessment of PCMM can have significant clinical implications for a small number of patients. Findings highlight the value of the SSMDT for high-quality care of patients with melanoma in the UK.
Elmore et al,[56] 2017	Consensus diagnosis by expert Interobserver	240 consecutive malignant melanoma cases	External-multiple institution Biopsy Surgical pathology		Concordance: Nevus/mild atypia: 92% (90-94%) Moderate atypia: 25% (22-28%) Severe atypia/ melanoma in-situ: 40% (37-44%) Early invasive melanoma pT1a: 43% (39-46%) Invasive melanoma (≥ pT1b): 72% (69-75%)	Diagnoses spanning moderately dysplastic nevi to early-stage invasive melanoma were neither reproducible nor accurate in this large study of pathologists in the USA. It is estimated that at a population level, 82.8% (81.0% to 84.5%) of melanocytic skin biopsy diagnoses would have their diagnosis verified if reviewed by a consensus reference panel of experienced pathologists, with 8.0% (6.2% to 9.9%) of cases over interpreted by the initial pathologist and 9.2% (8.8% to 9.6%) under interpreted.
Piepkorn et al,[57] 2019	Expert Retrospective	240 melanocytic lesions selected from a laboratory. Oversampled for difficult cases (dysplastic/spitz; melanoma in situ	External-multiple institution Biopsy Dermato- pathology		Mean agreement between physician pairwise rate for single interpretations of a case: 54.8%, κ = 0.42.	Second opinions rendered by dermato-pathologists improve reliability of melanocytic lesion diagnosis. However, discordance among pathologists remained high. Dermatopathologists are

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
		and thin melanoma; invasive thick melanoma) Strategies 1 – 6 based on initial interpretation of the case			Overall misclassification rate for a single interpretation [47.9% (46.7%-49.1%)] was the performance reference for second-opinion strategies 1-6: Strategy 1 (*second opinion on all skin biopsy samples*): 44.8% (42.5%- 47.1%; *P* < .001), the lowest misclassification rate Overinterpretation rates decreased from 6.3% (5.7%-6.9%) to 3.1% (2.4%-4.0%) Unchanged for under-interpretation at 41.6% (40.3%-43.0%) to 41.7% (39.2%-44.2%) Strategies 2 - 4 (*second opinions obtained when initial interpretations were melanoma in situ or invasive melanoma*) Overall misclassification rates similar to single interpretation Strategy 5 (*physician desire for second opinion*): 45.5% (43.5%–47.5%) Strategy 6 (*practice requires second opinion*): 45.3% (43.2%–47.4%) Misclassification rates for strategy 5 and 6 were decreased compared with the single interpretation rate: 47.9% (46.7%-49.1%, *P* < .001)	better at diagnosing melanocytic lesions than general pathologists.
Nethanel et al,[59]	Expert	177 consecutive malignant skin	External-single institution		Discordance rates Tis: 57%	A mandatory second review by expert is recommended for

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
2023	Retrospective	cases	Unknown/Not indicated Surgical pathology		T1b: 59% T3a: 67% T4a: 50% Alterations to the surgical plan: 15.3% of cases	Breslow 0.6-2.2mm as alterations could change surgical intervention. Incorporating such reviews into routine practice within dedicated cancer centers can improve diagnostic accuracy and guide appropriate treatment decisions, ultimately leading to better patient outcomes.
Lezcano et al,[47] 2023	Expert Retrospective	30 malignant skin cases, selection of cases unknown	Internal-multiple institution Unknown/Not indicated Surgical pathology		Pure vs mixed desmoplastic melanoma: κ = 0.64	Substantial agreement in the classification of the 30 tumors as pure versus mixed desmoplastic melanomas was found.
Gru et al,[12] 2018	Expert Prospective	75 Malignant, mycosis fungoides/Sézary syndrome cases, selection of cases unknown	External-multiple institution Unknown/Not indicated Surgical pathology, dermatopathology		Major discrepancies: 46.6% Minor discrepancies: 42.7%	Most of the major discrepancies were accounted by a difference in interpretation in the presence or absence of large cell transformation or FT. Most minor discrepancies were explained by a different interpretation in the expression of CD30. Digital slide scanning was beneficial, reliable, and practical for a methodical approach to perform central pathology review in the context of a large clinical prospective study.
Lohman et al,[31] 2021	Expert Retrospective	358 consecutive cases of malignant melanoma	External-single institution Biopsy Dermatopathology		Major discrepancies: 8.9%	There is value of dermatopathologic secondary review as a standardized part of a surgical referral.
Patrawala et al,[34] 2016	Expert Retrospective	588 cases of melanocytic lesions, all cases selected	External-single institution Unknown/Not indicated		Major discrepancies: 1.9%	Findings confirm consistent subjectivity in the histopathologic interpretation of

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
			Dermatopathology			melanoma. This study emphasizes that a review of the primary biopsy specimen may lead to significant changes in tumor classification, resulting in meaningful changes in clinical management.
Ronchi et al,[37] 2020	Expert Retrospective	121 consecutive cases of melanocytic lesions	External-single institution Biopsy Dermatopathology		Major discrepancies: 24.8%	Specialty dermatopathology opinion in the context of a multidisciplinary oncological team improves the diagnostic accuracy in selected difficult melanocytic lesions, significantly influencing the clinical choices and management of patients.
OTHER/MULTIPLE						
Bailey et al,[18] 2022	Expert Retrospective	928 consecutive consultation cases, type not reported	External-multiple institution Non-gyn cytology, FNAs Non-Gyn cytopathology		Major discrepancies (*a change in the discrete diagnostic category by two degrees or more*): 8.41% (78/928) Minor discrepancies: 28.45% (264/928)	Of the cases, 8.4% had major diagnostic discrepancies between the original diagnosis and the consultation diagnosis, which is consistent with reported values in surgical pathology consultation studies. The findings support the importance of second- opinion consultation in cytopathology to guide patient care.
Raju et al,[43] 2019	Unknown/Not indicated Retrospective	639 non-gyn and gyn cases, selection not indicated	Internal-single institution Biopsy, GYN cytology, non-GYN cytology Surgical pathology, both GYN and non-GYN cytopathology	Correlation of Pap test and gyn surgical pathology: 4.92% Correlation of non-gyn cytology and histo-pathology: 13.02%		The percentage of error in gynecologic cases was within the range of published data. However, those of non-gynecology cases were slightly higher. Regular professional training and improvement of sampling techniques especially the guided FNAC can reduce the errors. Regular professional training and improvement of

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
						sampling techniques especially the guided FNAC can reduce the errors. Hence, both clinicians and pathologists are equally responsible for the reduction of errors and improve patient care. Understanding the root cause and addressing it is the most important activity.
Kuijpers, et al,[29] 2016	Expert Retrospective	6796 multiple organ systems, clinician identified for MDC multi-disciplinary	Internal-single institution Unknown/Not indicated Surgical pathology		Total discordance: 34% (230/6796) Major discordance (*disagreement that affected patient care* - note: unclear who or how this assessment was made): 9% Minor discordance: 25% Change in diagnostic category: 3%	Benign cases had lowest discordance; followed by malignant cases, where the majority of discordances were 'minor'. Undetermined malignant potential had the highest percentage of major discordance. Routine double reading of histopathology specimens prior to multidisciplinary meetings prevents diagnostic errors. It resulted in about 1% discordant diagnoses of potential clinical significance, indicating that second review is worthwhile in terms of patient safety and quality of patient care.
Peck et al,[35] 2018	Pathologist without designated expertise Retrospective	187 mixed cases, selection of cases unknown	External-multiple institution Unknown/Not indicated Surgical pathology		Major Discordance Malakoplakia 22% Herpetic ulcer 23% Leucocytoclastic vasculitis 3% Adenocarcinoma, diffuse, Helicobacter gastritis 73% Spirochaetosis 17% Secretory endometrium 22% Necrotising sialometaplasia-like changes 23%	Data indicated potentially serious diagnostic inaccuracy in histopathology; however, this may not reflect real-world numbers and the findings of the 2015 statement are reiterated. Second opinion is 'effective' and 'integral' to quality for regional labs. Clinical/ radiologic findings are critical to accurate diagnosis.

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
					HNC - metastatic SCC & metastatic medullary thyroid carcinoma 18% Rheumatoid nodule 2% Minor Discordance Malakoplakia 4% Herpetic ulcer 0% Leucocytoclastic vasculitis 30% Adenocarcinoma, diffuse, Helicobacter gastritis 0% Spirochaetosis 32% Secretory endometrium 8% Necrotising sialometaplasia-like changes 5% HNC - metastatic SCC & metastatic medullary thyroid carcinoma 4% Rheumatoid nodule 40%	
Vats et al,[40] 2022	Expert Retrospective	457 unknown/ not specified cases	External-multiple institution Biopsy, excisional biopsy, resection Surgical pathology		Malignant to benign: 20 cases (4.4%) Benign to malignant: 10 cases (2.2%) Major: 6.56% (30/457)	Smaller practices benefit from referral centers that include pathologists, with additional training in the diagnosis and management of STS, as well as access to molecular and extended IHC techniques necessary for accurate diagnosis.
Middleton et al,[42] 2016	Expert Retrospective	13109 unknown/not specified cases	External-multiple institution Unknown/Not indicated Not stated		All discrepancies: 11.35% (1488/13109) Major (*diagnosis changes treatment or surveillance*): 2.74% (359/13109) Minor: 8.61% (1129/13109)	Diagnostic errors may cause harm to patients by preventing or delaying treatment. Secondary review can significantly improve outcomes.
Solivas et al,[44] 2024	Another board-certified pathologist within the same group	All surgical pathology cases at one institution which underwent review from 2015	Internal-single institution Biopsy, FNA Surgical pathology		96.9% concordance 3.1% discordance Rate of diagnostic change 2.0%.	The low rate of diagnostic change in the studied institution might be attributed to good diagnostic accuracy. However, it is also possible

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
	Prospective	to 2018, N=5377				that reviewing pathologists tended to agree with the diagnosis made by their colleagues because of the element of peer pressure.
Nagose et al,[46] 2022	Expert Retrospective	60 serous effusion, selection of cases unknown	Internal-single institution Not indicated Non-GYN cytopathology		All participants vs standard: k=0.212 Between participants: k=0.327 The range of κ for interobserver agreement of fellows: 0.169–0.725	The participant with substantial diagnostic agreement with expert (κ = 0.729, 88.3%) had the most experience. Causes of major discordance were pertaining to paucity and distribution of cells, and to misinterpretation of reactive mesothelial cells.
Rakheja et al,[51] 2022	Pathologist without designated expertise Retrospective	316 FNA of soft tissue tumors	Internal-single institution FNA Surgical pathology, non-GYN cytopathology	Sensitivity: 95.3% Specificity: 99.6% PPV: 97.6% NPV: 99.2% Accuracy: 99.0%		With aid of ancillary cellblock preparation and immune-cytochemistry, FNAC can serve as invaluable tool, for not only primary lesions, but also for metastatic tumors as well as documentation of locally recurrent soft tissue neoplasms.
Shinohara et al,[52] 2022	Expert Retrospective	270 consecutive kidney cases	External-multiple institution Not indicated Surgical pathology		Original inconclusive diagnosis was changed for 52% of the total diagnostically difficult cases. Minor disagreement to the original diagnosis: 12% Major disagreement to the original diagnosis: 10% Major discrepancy (*change diagnosis to require different treatments*) 12%	The impact of WSI-based expert consultation on inconclusive histopathological diagnoses was studied and showed substantial improvement in diagnosis after WSI-based remote second opinion.
Torous et al,[53] 2022	Expert Retrospective	202 cases amended cases with diagnostic error, selected for consensus diagnosis	Internal-single institution (review of cases diagnosed at one institution) FNA		Specimen Identification Error GYN 15.7% Non-GYN 12.6% General Report Defects GYN 68.7% Non-GYN 25.2%	Studying amendment reports is an underrecognized and valuable quality assurance tool. Amendments can help provide information about types of errors, monitor laboratory processes, and help guide

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
			GYN surgical pathology, Non-GYN cytopathology		Diagnostic Error GYN 15.7% Non-GYN 62.2%	quality improvement endeavors.
Gupta et al,[54] 2022	Expert Retrospective	151 consecutive kidney cases of eyeballing vs digital measurement	Internal-single institution Biopsy Surgical pathology	Grade 0 Sensitivity: 72.45% Specificity: 98.11% PPV: 98.61% NPV: 65.82% Accuracy: 81.40% Grade 1 Sensitivity: 52% Specificity: 78.57% PPV: 32.50% NPV: 89.19% Accuracy: 74.17% Grade 2 Sensitivity: 47.83% Specificity: 89.06% PPV: 44% NPV: 90.48% Accuracy: 82.78% Grade 3 Sensitivity: 100% Specificity: 93.84% PPV: 95.71% NPV: 100% Accuracy: 94.03%	Concordance between "eyeballing" and digital measurement of grade overall: 66.22%	Assessment of tubular atrophy on digital images will be the way forward for accurate quantification.
Wellauer et al,[55] 2022	Expert Retrospective	347 consecutive sarcoma cases	Both internal and external samples of institution A were initially analyzed by the local pathology institute. Afterwards, the samples being reviewed and assessed by a reference institute pathologist specialized in soft tissue tumors (external). Conversely,		Major discrepancies (*influenced the treatment strategy directly*): 12.2% Minor discrepancies: 10.2% Diagnostic concordance: 77.6%	Results confirm the importance of a pathological second review by a reference pathologist.

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity Positive Predictive Value (PPV) Negative Predictive Value (NPV) Accuracy Other (95% Confidence Interval)	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
			institution B cases were assessed directly by the reference institute pathologist (internal). Biopsy Surgical pathology			
Giunchi et al,[27] 2017	Expert Retrospective	116 prostate lesions for consensus diagnosis	External-multiple institution Biopsy, excisional biopsy, resection Surgical pathology		Overall agreement and kappa between all possible pairs of reviewers: Prostate cancer: 80.2%, k=0.84 PIN: 67.2%, k=0.42 Proliferative atrophic lesions: 49.1%, k=0.32 When using as gold standard the assessment of a single genitourinary pathologist, mean agreement: Prostate cancer: 96.9%, k=0.88 PIN: 92.9%, k=0.55 Proliferative atrophic lesions: 72.2%, k=0.55	The study data show differences in the agreement with the gold standard for specialist genitourinary and general European pathologists were minimal.

Abbreviations: AUC, area under the curve; AUS, atypia of undetermined significance; BX, biopsy; CI, confidence interval; DCIS, ductal carcinoma in situ; ENT, ear, nose, and throat; ER, estrogen receptor; FISH, fluorescence in situ hybridization; FLUS, follicular lesion of undetermined significance; FNA, fine-needle aspiration; FNAB, fine-needle aspiration biopsy; FN/SFN, follicular neoplasm/suspicious for follicular neoplasm; GCC, goblet cell carcinoid; GI, gastrointestinal; GYN, gynecology; HER2, human epidermal growth factor receptor 2; HGD, high-grade dysplasia; HPF, high power field; IAC, International Academy of Cytology; IHC, immunohistochemistry; IG, intermediate grade; IND, indeterminate; Ki67, antigen Kiel 67; LAMN, low-grade appendiceal mucinous neoplasm; LG, low grade; LGD, low-grade dysplasia; MDC, multidisciplinary care; NPV, negative predictive value; NSCLC, non-small cell lung cancer; PCMM, primary cutaneous malignant melanoma; PgR, progesterone receptor; PIN, prostatic intraepithelial neoplasia; PPV, positive predictive value; ROSE, rapid on-site evaluation; ROM, risk of malignancy; SHCN, suspicious for Hurthle cell neoplasm; SN, sentinel node; SN-positive/SN-negative, sentinel node positive/negative; SOD, second-opinion diagnosis; SSMDT, specialist skin multidisciplinary team; TBSRTC, The Bethesda System for Reporting Thyroid Cytopathology; TIC, touch imprint cytology; UNSAT, unsatisfactory; WSI, whole slide imaging.

Supplemental Table 9. Case Reviews

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity PPV, NPV Accuracy Other	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
BREAST						
Agrawal et al,[16] 2021	Pathologist without designated expertise Retrospective	1147 consecutive breast cases	Internal-single institution Non-gyn cytology Surgical pathology, non-Gyn cytopathology	Malignant cytopathology case found malignant on histopathology With ROSE Sensitivity: 92.9% Specificity: 100% PPV: 100% NPV: 94.9% Accuracy: 96.9% AUC: 0.97 Without ROSE Sensitivity: 82.1% Specificity: 99.4% PPV: 99.2% NPV: 86.9% Accuracy: 91.5% AUC: 0.91 Suspicious and malignant cytopathology case found in situ or malignant on histopathology With ROSE Sensitivity: 98.8% Specificity: 99.1% PPV: 98.8% NPV: 99.1% Accuracy: 99.0% AUC: 0.99 Without ROSE Sensitivity: 99.3% Specificity: 99.4% PPV: 99.3% NPV: 99.4% Accuracy: 99.4% AUC: 0.99%		Incorporating the IAC Yokohama System for Reporting Breast FNAB Cytopathology with ROSE can improve early and accurate diagnosis of breast lesions, prevent missed diagnoses, and provide reliable estimates of ROM in a given patient population.
Elmore et al,[71] 2015	Expert Retrospective	6900 randomly selected benign, atypical, DCIS, IDC cases	External-multiple institution Excisional biopsy Surgical pathology		Overinterpretation Benign: 13% Atypical: 17% DCIS: 3%	Results show that atypia is a diagnostic classification with considerable variation among practicing pathologists, with an overall concordance rate of 48%

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity PPV, NPV Accuracy Other	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
					Under-interpretation Atypical: 35% DCIS: 13% IDC: 4%	compared with the consensus-derived reference diagnosis. Pathologists from outside of academic settings, those who interpret lower weekly volumes of breast cases and those from small-sized practices were statistically significantly less likely to agree with the consensus-derived reference diagnosis.
Hohnen et al,[67] 2021	Expert Retrospective	2036 mixed breast cases, selection not indicated	Internal-multiple institution Biopsy Surgical pathology		Benign: 56.04% (1141/2036) Malignant: 34.33% (699/2036) Intermediate: 7.22% (147/2036) Non-diagnostic: 2.26% (46/2036) Discrepancies: 49% All cases of major discordant disagreements would have resulted in significant changes to clinical management.	This study demonstrates that an Australian institution is providing a high-quality pathology service with a low error rate between initial and review diagnoses of breast core biopsies. It reinforces the importance of secondary review of biopsies in a timely fashion for detecting potentially serious misdiagnoses that could lead to inappropriate management.
Khazai et al,[28] 2015	Expert Retrospective	1970 breast cases referred for cancer	External-multiple institution Unknown/Not indicated Surgical pathology		Significant discrepancies: 11.47% (226/1970)	A 11.47% significant error rate for cases referred to the Breast service at authors' institution in 2010 was calculated.
Woeste et al,[50] 2022	Expert Retrospective	263 consecutive breast cases	External-single institution Excisional biopsy Surgical pathology		Discordant diagnosis 35% Change in tumor subtype 10% Change in invasive cancer grade 45%	Although routine expert breast pathology consultation resulted in few changes in clinical and operative management within this cohort, it ultimately served its purpose as a quality assurance measure ensuring that each patient received the correct treatment. Excisional biopsies and benign lesions without atypia may be associated with less diagnostic discordance. However, more studies are needed to evaluate risk

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity PPV, NPV Accuracy Other	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
						for pathologic discordance after expert pathology consultation for a case-by-case approach to be accepted as an adequate substitute. Expert pathology review remains essential in the multidisciplinary care of all breast cancer patients.
ENDOCRINE						
Agrawal et al,[62] 2020	Pathologist without designated expertise Retrospective	65 thymoma consecutive cases	Internal-single institution Resection Surgical pathology		K = 0.73	Familiarity with the WHO thymoma criteria helps to improve interobserver variability in assigning categories. Especially in the B category of thymomas.
Bellevicine et al,[21] 2017	Expert Retrospective	34 consecutive thyroid cases	External-multiple institution FNAs Non-GYN cytopathology		Concordance: 64.71% (22/65) Discordance: 35%	Intra-institutional second opinion practice for "indeterminate" thyroid FNA avoids unnecessary surgeries and maximises the detection of malignant cases diagnosed as FN/SFN.
Vanden Bussche et al,[73] 2015	Expert Retrospective	272 atypical/ suspicious thyroid cases	Internal-single institution FNAs Non-GYN cytopathology		κ: $P < .001$ AUS-HC: 60% SHCN: 51%	Hürthle cells are arguably the most common cause of interobserver disagreement rates for the interpretation of indeterminate nodules. One-step differences, especially SHCN to/from AUS-HC, appear to be common.
Sripodok et al,[49] 2023	Pathologist without designated expertise Retrospective	All case with borderline, atypia of undetermined significance/FLUS diagnosis and follow up histology, N=72	Internal-single institution FNA Non-GYN cytopathology		Consensus diagnosis and pathologist interpretation κ range 0.674 to 0.898	There are significant interobserver variations, and changes in cytological diagnosis when performing re-evaluation among different pathologists using TBSRTC. A second or third opinion should be routinely sought to establish a consensus diagnosis as an adjunctive method for an initial AUS/FLUS diagnosis. The reclassification lowers medical costs and the rate of unnecessary surgery, particularly in patients with cytologically confirmed benign thyroid nodules. Under the condition that the cytological sample is

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity PPV, NPV Accuracy Other	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
						cellularly adequate, preoperative molecular evaluation is a promising method for aiding in the diagnosis of thyroid nodules that requires additional study.
ENT/HEAD & NECK						
Adhya et al,[9] 2019	Unknown/Not indicated Prospective	280 consecutive malignant oral cavity cases	Internal-single institution Non-GYN cytology Surgical pathology, non-GYN cytopathology	Sen: 0.982 Spec: 0.893 PPV: 0.973 NPV: 0.926 Accuracy: 0.964		TIC aids in the diagnosis of oral squamous cell carcinoma but does not replace incisional biopsy. The cytologic examination of imprints of biopsies also helps determine the adequacy of the specimen for histologic examination.
Richards et al,[48] 2023	Expert Retrospective	23 consecutive borderline, oral verrucous squamous lesion cases	Internal-single institution Biopsy Surgical pathology		Diagnostic concordance between biopsy and resection (with respect to a definitive diagnosis on biopsy and agreement on resection): 17%	
GASTROINTESTINAL						
Van Der Wel et al,[58] 2019	Mixture of experts and non-experts Retrospective	55 cases of Barrett's	External-multiple institution Biopsy Surgical pathology		NDBO Discordance 21.2% Major over interpretation: HGD: 28.9% LGD: 58.4% IND: 77.1% Major diagnostic error in 8.80% of all diagnoses	Data provide evidence-based criteria for diagnostic proficiency in Barrett's histopathology
Govind et al,[72] 2020	Pathologist without designated expertise Prospective	50 cases of gastric neuroendocrine tumor, selection not indicated	Internal-multiple institution Unknown/Not indicated Surgical pathology	Training accuracy: 98.4%, Validation accuracy: 90.9% Testing accuracy: 91.0%		The proposed methods can improve accuracy and potentially save a significant amount of time if implemented into clinical practice.
GENITOURINARY						
Abro et al,[61] 2021	Expert Retrospective	106 consecutive bladder cases	Internal-single institution Biopsy, non-GYN cytology, bladder washing	Determining NPV for UNSAT urine cytology: 84.4% For bladder washing specimens, the NPV for UNSAT was 87.5%, versus		In a cohort of patients sampled with various UCyto specimen types obtained for UC surveillance and for other indications, the NPV of an UNSAT diagnosis was 90.0%

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity PPV, NPV Accuracy Other	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
			Surgical pathology, non-GYN cytopathology	96.2% for "normal" results, whereas for other specimen types, including voided urine specimens, they were 75.0% versus 98.4%, respectively.		
Gordetsky et al,[66] 2018	Expert Retrospective	246 consecutive malignant bladder cases	External-multiple institution Biopsy, excisional biopsy, resection Surgical pathology		Any change in diagnosis: 36.99% (91/246) Major change in diagnosis: 19.11% (47/246) Change in malignancy status: 8.79% (8/91) Change in stage: 50.5% Change in tumor type: 17.6% Change in histologic subtype of urothelial carcinoma: 17.6% Change in grade:15.4%	Secondary pathology review of bladder specimens by a GU subspecialized surgical pathologist can show a major change in diagnosis in almost 20% of cases.
GYNECOLOGY						
Banet et al,[20] 2015	Expert Retrospective	237 consecutive pelvic washings cases	External-multiple institution GYN cytology GYN cytopathology		Major discordance (*counted if the diagnosis changed from negative to positive, or vice versa*): 3.80% (9/237) Minor discordance: 12.24% (29/237)	Only a few cases showed major or minor discordance, and of those that were discordant, most involved ovarian specimens. Major discordance between original and consulting diagnoses would have changed treatment in only 2 cases.
LUNG						
Witt et al,[74] 2016	Pathologist without designated expertise Retrospective	50 lung cancer - adeno vs SCC cases, selection not indicated	Internal-single institution Biopsy, FNAs Surgical pathology; non-GYN cytopathology	The overall accuracy for differentiating between SCC and ADC of the lung was 65% (95% CI 58-72%) with the indeterminate category and 66% (95% CI 60-72%) without.	The inter-observer agreement With in-determinate category: 0.22 Without in-determinate category: 0.1	Overall, the low interobserver agreement in the study indicates that accurate subclassification between the NSCLCs often cannot be made by cytomorphology alone.
SKIN						
Balasub	Expert	3006 consecutive	Internal-single institution		Concordant: 59.81%	Several deficits were identified that

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity PPV, NPV Accuracy Other	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
ramanian et al,[19] 2015	Retrospective	skin cases	Biopsy Surgical Pathology		(1798/3006) Partially concordant: 7.58% (228/3006) Discordant: 30.90% (929/3006)	need to be rectified to improve the diagnostic accuracy of skin biopsy. A portion of discordant slides showing features compatible with the disease when reviewed by a pathologist and dermatologist together emphasizes the importance of a joint review by both in doubtful cases.
Bhoyrul et al,[22] 2019	Expert Retrospective	341 consecutive melanoma cases	External-multiple institution Biopsy, Excisional biopsy Surgical pathology		Inter-observer variability 6.74% (23/341) Change management 2.93% (10/341)	Disagreements in the pathological assessment of PCMM can have significant clinical implications for a small number of patients. Findings highlight the value of the SSMDT for high-quality care of patients with melanoma in the UK.
Cocks et al,[64] 2021	Expert Retrospective	7950 amended reports - multi-institutional submission of cases evaluated at a central institution by expert pathologists	Internal-single institution Biopsy, Excisional biopsy Surgical pathology		Major diagnostic changes (*potential to alter patient care*): 27.9%	Highlights quality improvement areas that could be targeted to avoid diagnostic errors especially seeking clinics-pathological correlation in challenging melanocytic lesions.
El Sharouni et al,[65] 2021	Expert Retrospective	322 consecutive sentinel lymph nodes for melanoma cases	External-multiple institution Biopsy, Excisional biopsy Surgical pathology		Positive for metastasis to negative on review: 11.80% (38/322) 4.3% of patients with false positive diagnoses for metastatic melanoma would have qualified for adjuvant therapy (*unjustified based on false positive diagnoses after expert review*) The 5-year RFS of the 38 downgraded patients was 86.7%, similar to the 85.9% for 6413 SN-negative patients and better than the 53.2% of 284 patients who were truly SN positive upon review.	A large number of originally positive SN biopsies in patients with melanoma are misclassified, indicating that some patients with melanoma might receive unjustified adjuvant treatment. We therefore advocate that when adjuvant treatment is considered in patients with stage III melanoma, SN biopsies should be reassessed by an expert melanoma pathologist.

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity PPV, NPV Accuracy Other	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
OTHER/MULTIPLE						
Ahmad et al,[63] 2016	Unknown/Not indicated Not applicable/ Not indicated	199 cases at clinician request, type not specified	Internal-single institution Surgical pathology		Errors identified in paraffin embedded specimen reports, N=199 Misinterpretation due to failure to perform appropriate IHC stains 12.10% Misinterpretation due to failure to submit appropriate/adequate sections for histologic examination 12.60% Misinterpretation per se (*i.e. not due to lack of any appropriate workup*) 15.10% Missing critical/important information in reports 32.70% Distribution of misinterpretations in various organs (n=99) GI tract 23.20% Breast 13.10% Lungs,pleura, mediastinum 10.10% Lymph nodes 9.10% Head & Neck 9.10% Kidney &urinary bladder 8.10% Female genital tract 8.10% Bone and Soft tissue 7.10% Brain 6.10% Skin 4.00% Male Genital Tract 2.00%	Error rate higher when case review is requested by clinicians
Bailey et	Expert	928 consecutive	External-multiple institution		Major discrepancies (*a*	Of the cases, 8.4% had major

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity PPV, NPV Accuracy Other	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
al,[18] 2022	Retrospective	consultation cases, type not reported	Non-gyn cytology, FNAs Non-Gyn cytopathology		*change in the discrete diagnostic category by two degrees or more*): 8.41% (78/928) Minor discrepancies: 28.45% (264/928)	diagnostic discrepancies between the original diagnosis and the consultation diagnosis, which is consistent with reported values in surgical pathology consultation studies. The findings support the importance of second- opinion consultation in cytopathology to guide patient care.
Raju et al,[43]	Unknown/Not indicated Retrospective	639 non-gyn and gyn cases, selection not indicated	Internal-single institution Biopsy, GYN cytology, non-GYN cytology Surgical pathology, both GYN and non-GYN cytopathology	Correlation of Pap test and gyn surgical pathology: 4.92% Correlation of non-gyn cytology and histo-pathology: 13.02%		The percentage of error in gynecologic cases was within the range of published data. However, those of non-gynecology cases were slightly higher. Regular professional training and improvement of sampling techniques especially the guided FNAC can reduce the errors. Regular professional training and improvement of sampling techniques especially the guided FNAC can reduce the errors. Hence, both clinicians and pathologists are equally responsible for the reduction of errors and improve patient care. Understanding the root cause and addressing it is the most important activity.
Johnson et al,[68] 2021	Expert Retrospective	740 surgical cases sent to institution for referral within a designated time frame	Internal-multiple institution Biopsy Surgical pathology		Diagnostic discrepancies were identified in 104 (14.1%) patients, 30 (4.1%) of which resulted in a change in care.	Preventable diagnostic errors are reduced by pathology review for patients referred within a single health care system.
Kuijpers et al,[29] 2016	Expert Retrospective	6796 multiple organ systems, clinician identified for MDC multi-disciplinary	Internal-single institution Unknown/Not indicated Surgical pathology		Total discordance: 34% (230/6796) Major discordance (*disagreement that affected patient care* - note: unclear who or how this assessment was made): 9% Minor discordance: 25% Change in diagnostic	Benign cases had lowest discordance; followed by malignant cases, where the majority of discordances were 'minor'. Undetermined malignant potential had the highest percentage of major discordance. Routine double reading of histopathology specimens prior to multidisciplinary meetings prevents diagnostic errors. It resulted in

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity PPV, NPV Accuracy Other	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
					category: 3%	about 1% discordant diagnoses of potential clinical significance, indicating that second review is worthwhile in terms of patient safety and quality of patient care.
Kuijpers et al,[60] 2015	Expert - Specialized cyto-pathologists routinely re-diagnosed blinded defined categories of clinical cytology specimens that had been signed out by routine pathologists Prospective	218 multiple organ systems cases selected for consensus diagnosis	Internal-single institution All cytology samples Non-GYN cytopathology		Discrepancies: 12.84% (28/218) Initial and expert diagnoses were concordant in 131/218 specimens (60.1 %). Major and minor discordances were present in 28 (12.8 %) and 59 (27.1 %) specimens, respectively. Pleural fluid, thyroid and urine specimens showed the highest major discordance rates (19.4, 19.2 and 16.7 %, respectively). Histological follow-up (where possible) supported the expert diagnosis in 95.5 % of specimens. The implemented double reading strategy of defined categories of cytology specimens showed major discordance in 12.8 % of specimens. The expert diagnosis was supported in 95.5 % of discordant cases where histological follow-up was available.	Results emphasize that cytopathology is a subspecialization of pathology and requires specialized cytopathologists. This indicates that this double reading strategy is worthwhile and contributes to better cytodiagnostics and quality of patient care, especially for suspicious pleural fluid, thyroid and urine specimens.
Layfield et al,[69] 2017	Expert review for all pathology review protocols.	2466 random review cancer cases, 70 solicited cases, 190 unsolicited cases by outside institutions, 59 dermato-pathology cases	Internal-single institution Unknown/Not indicated Surgical pathology		The 10% random review detected 17 errors in 2147 cases (0.79%). Solicited case consultations requested by clinicians or internal pathologists detected 5 diagnostic errors in 70 cases (7.1%). Unsolicited reviews by outside institutions in the	Focused reviews initiated by diagnostic concerns of a clinician or pathologist, unsolicited reviews because of treatment at another institution and sub-specialty-based reviews appear to be more effective in detecting diagnostic errors than the 10% random review. Quality assurance programs should include focused reviews in addition to 10% random review to maximize error detection

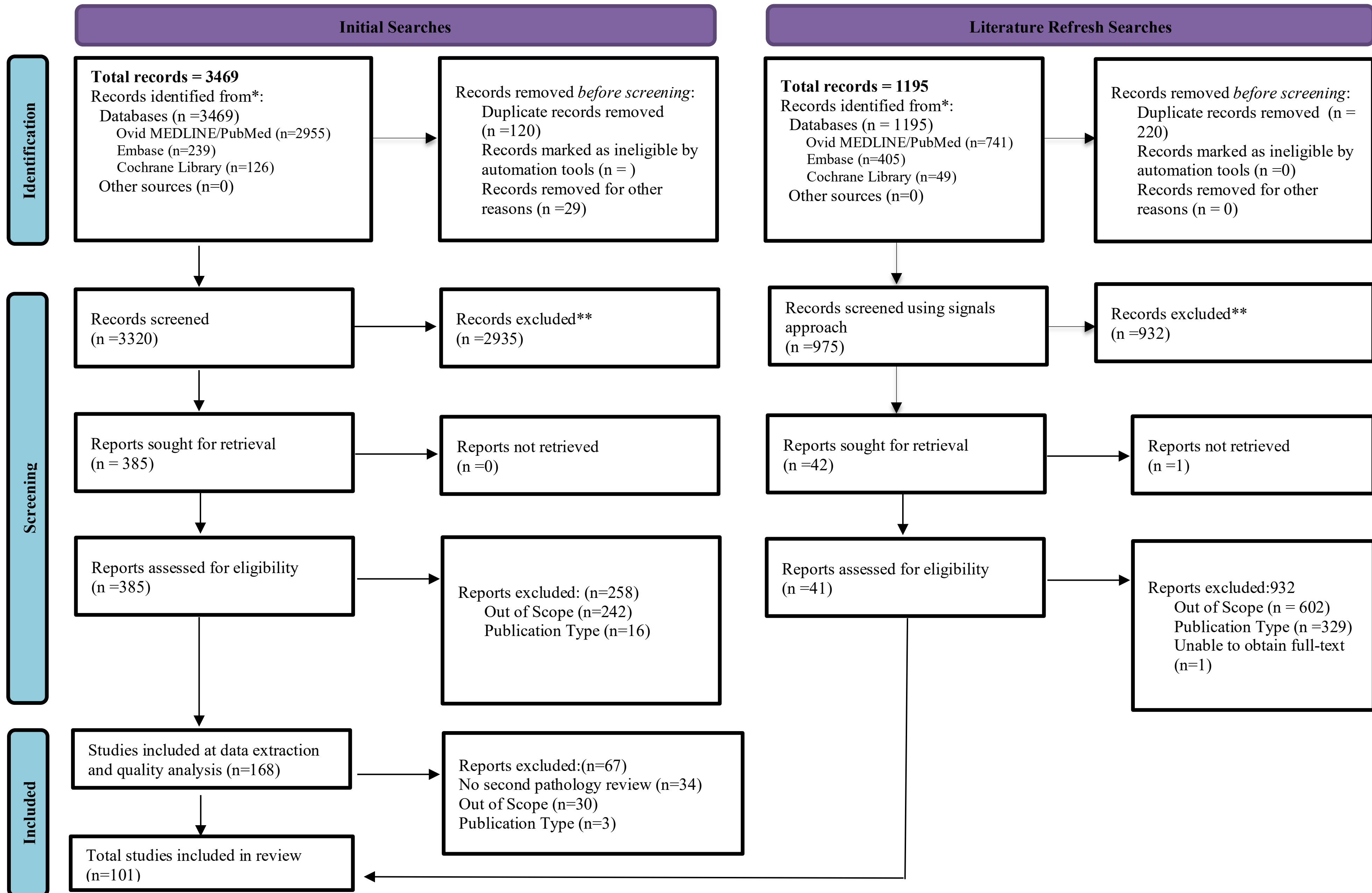
Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity PPV, NPV Accuracy Other	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
					course of patient care detected 3 diagnostic errors in 190 cases (1.6%). Review of the dermatopathology material disclosed 5 diagnostic errors in 59 cases (8.5%).	
Middleton et al,[42] 2016	Expert Retrospective	13109 unknown/not specified cases	External-multiple institution Unknown/Not indicated Not stated		All discrepancies: 11.35% (1488/13109) Major (*diagnosis changes treatment or surveillance*): 2.74% (359/13109) Minor: 8.61% (1129/13109)	Diagnostic errors may cause harm to patients by preventing or delaying treatment. Secondary review can significantly improve outcomes.
Yue et al,[70] 2020	Expert Retrospective	457 infection cases, selection not specified	External-multiple institution Unknown/Not indicated Surgical pathology		Overall error: 28.01% (128/457) Clinically significant difference: 20.6% Non-clinically significant difference: 7.44% (34/457) Nature of clinically significant discrepancy between original report and specialist review Samples (total=94) Incorrectly suspected infection 48.9% Missed infection 20.2% Different infection 8.5% More specific identification of organism/viability 21.3%	A review of histopathology cases by an infectious disease (ID) histopathology referral centre has yielded a 20.6% clinically significant error rate. Measures to improve training in ID histopathology in the UK are discussed.
Rakheja et al,[51] 2022	Pathologist without designated expertise Retrospective	316 soft tissue tumor cases	Internal-single institution FNA Surgical pathology, non-GYN cytopathology	Sensitivity: 95.3% Specificity: 99.6% PPV: 97.6% NPV: 99.2% Accuracy: 99.0%		With aid of ancillary cellblock preparation and immunocytochemistry, FNAC can serve as invaluable tool, for not only primary lesions, but also for metastatic tumors as well as

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity PPV, NPV Accuracy Other	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
						documentation of locally recurrent soft tissue neoplasms.
Shinohara et al,[52] 2022	Expert Retrospective	270 consecutive kidney cases	External-multiple institution Not indicated Surgical pathology		Original inconclusive diagnosis was changed for 52% of the total diagnostically difficult cases. Minor disagreement to the original diagnosis: 12% Major disagreement to the original diagnosis: 10% Major discrepancy (*change diagnosis to require different treatments*) 12%	The impact of WSI-based expert consultation on inconclusive histopathological diagnoses was studied and showed substantial improvement in diagnosis after WSI-based remote second opinion.
Torous et al,[53] 2022	Expert Retrospective	202 cases amended cases with diagnostic error, selected for consensus diagnosis	Internal-single institution (review of cases diagnosed at one institution) FNA GYN surgical pathology, non-GYN cytopathology		Specimen Identification Error GYN 15.7% Non-GYN 12.6% General Report Defects GYN 68.7% Non-GYN 25.2% Diagnostic Error GYN 15.7% Non-GYN 62.2%	Studying amendment reports is an underrecognized and valuable quality assurance tool. Amendments can help provide information about types of errors, monitor laboratory processes, and help guide quality improvement endeavors.
Gupta et al,[54] 2022	Expert Retrospective	151 consecutive kidney cases, eyeballing vs digital measurement	Internal-single institution Biopsy Surgical pathology	Grade 0 Sensitivity: 72.45% Specificity: 98.11% PPV: 98.61% NPV: 65.82% Accuracy: 81.40% Grade 1 Sensitivity: 52% Specificity: 78.57% PPV: 32.50% NPV: 89.19% Accuracy: 74.17% Grade 2 Sensitivity: 47.83% Specificity: 89.06% PPV: 44%	Concordance between "eyeballing" and digital measurement of grade overall: 66.22%	Assessment of tubular atrophy on digital images will be the way forward for accurate quantification.

Study	Expert Review Type	Number of Cases and Case selection	Setting Procedure type Discipline	Sensitivity Specificity PPV, NPV Accuracy Other	Kappa (κ) Concordance/ Discordance Discrepancies (95% Confidence Interval)	Authors' Conclusions
				NPV: 90.48% Accuracy: 82.78% Grade 3 Sensitivity: 100% Specificity: 93.84% PPV: 95.71% NPV: 100% Accuracy: 94.03%		
Wellauer et al,[55] 2022	Expert Retrospective	347 consecutive sarcoma cases	Both internal and external samples of institution A were initially analyzed by the local pathology institute. Afterwards, the samples being reviewed and assessed by a reference institute pathologist specialized in soft tissue tumors (external). Conversely, institution B cases were assessed directly by the reference institute pathologist (internal). Biopsy Surgical pathology		Major discrepancies (*influenced the treatment strategy directly*): 12.2% Minor discrepancies: 10.2% Diagnostic concordance: 77.6%	Results confirm the importance of a pathological second review by a reference pathologist.

Abbreviations: ADC, adenocarcinoma; AUS, atypia of undetermined significance; BX, biopsy; CI, confidence interval; DCIS, ductal carcinoma in situ; ER, estrogen receptor; FISH, fluorescence in situ hybridization; FNA, fine-needle aspiration; GYN, gynecology; HER2, human epidermal growth factor receptor 2; IDC, invasive ductal carcinoma; IHC, immunohistochemistry; Ki67, antigen Kiel 67; LG, low grade; MDC, multidisciplinary care; NPV, negative predictive value; NSCLC, non-small cell lung cancer; PCMM, primary cutaneous malignant melanoma; PgR, progesterone receptor; PPV, positive predictive value; RFNA, repeat fine-needle aspiration; RFS, recurrence-free survival; ROM, risk of malignancy; ROSE, rapid on-site evaluation; SHCN, suspicious for Hurthle cell neoplasm; SN, sentinel node; SOD, second-opinion diagnosis; SSMDT, specialist skin multidisciplinary team; TBSRTC, The Bethesda System for Reporting Thyroid Cytopathology; UCyto, urinary cytology; WHO, World Health Organization; WSI, whole slide imaging.

Supplemental Figure 1: Systematic Literature Review Flow Diagram



Adapted From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: http://www.prisma-statement.org/

Supplemental Figure 2: Database Search Strings

Ovid MEDLINE Search String:

1 exp Diagnostic Errors/
2 delayed diagnosis/ or incidental findings/ or overdiagnosis/
3 Near Miss, Healthcare/
4 Incidental Findings/
5 (over$diagno$ or miss$diagnos$ or mis$diagnos$ or misdiagnos$).tw.
6 (diagnos$ adj1 (miss or missed or mis or over)).tw.
7 (error$ adj2 (reduc$ or omission$ or commission$ or communication$ or consultation$ or critical)).tw.
8 "near event$".tw.
9 "near miss$".tw.
10 (event$ adj1 (sentinal or adverse or critical)).tw.
11 (report$ adj1 amend$).tw.
12 "second opinion$".tw.
13 (diagno$ adj2 (error$ or delay$ or accuracy or uncertain$)).tw.
14 ((intra$operative or inter$institutional) adj1 (consult$ or communication or communications)).tw.
15 ((observer or interobserver) adj1 (variation$ or variability or agreements or agreement or consultat$ or reproduc$)).tw.
16 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15
17 exp Reproducibility of Results/
18 (quality adj1 (assurance or improv$)).tw.
19 "discordant rate$".tw.
20 miss rate$.tw.
21 (error$ adj1 (rate$ or reduct$ or detect$)).tw.
22 (imprecision$ or concordance or discordance or reproducibility or discrepan$ or ambigu$ or indeterm$).tw.
23 17 or 18 or 19 or 20 or 21 or 22
24 Pathology, Surgical/
25 exp Histocytochemistry/
26 exp Cytodiagnosis/
27 exp Histocytological Preparation Techniques/
28 Pathology, Molecular/
29 exp cytological techniques/ or exp histological techniques/ or molecular diagnostic techniques/
30 ((anatomic$ or molecular or histo$ or surgical) adj pathology).tw.
31 (cytopath$ or histopath$ or immunohisto$ or histocyto$ or immunocyto$).tw.
32 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31
33 16 and 23 and 32
34 limit 33 to (English language and yr="2015 -Current")
35 limit 34 to (comment or editorial or letter or news or newspaper article or observational study, veterinary or personal narrative)
36 34 not 35

Embase Search String:

1. 'Diagnostic Errors'/exp
2. 'delayed diagnosis'/de OR 'incidental findings'/de OR overdiagnosis/de
3. 'Near Miss, Healthcare'/de
4. 'incidental findings'/de
5. (over?diagno? OR miss?diagnos? OR mis?diagnos? OR misdiagnos?)
6. (diagnos? NEAR/1 (miss OR missed OR mis OR over))
7. error? NEAR/2 (reduc? OR omission? OR commission? OR communication? OR consultation? OR critical)
8. 'near event?'
9. 'near miss?'
10. (event? NEAR/1 (sentinal OR adverse OR critical))
11. (report? NEAR/1 amend?)
12. 'second opinion?'
13. (diagno? NEAR/2 (error? OR delay? OR accuracy OR uncertain?))
14. ((intra?operative OR inter?institutional) NEAR/1 (consult? OR communication OR communications))
15. ((observer OR interobserver) NEAR/1 (variation? OR variability OR agreements OR agreement OR consultat? OR reproduc?))
16. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15
17. 'Reproducibility of Results'/exp
18. (quality NEAR/1 (assurance OR improv?))
19. 'discordant rate?'
20. 'miss rate?'
21. (error? NEAR/1 (rate? OR reduct? OR detect?))
22. (imprecision? OR concordance OR discordance OR reproducibility OR discrepan? OR ambigu? OR indeterm?)
23. #17 OR #18 OR #19 OR #20 OR #21 OR #22
24. 'Pathology, Surgical'/de
25. Histocytochemistry/exp
26. Cytodiagnosis/exp
27. 'Histocytological Preparation Techniques'/exp
28. 'Pathology, Molecular'/de
29. 'cytological techniques'/exp OR 'histological techniques'/exp OR 'molecular diagnostic techniques'/de
30. ((anatomic? OR molecular OR histo? OR surgical) NEXT/1 pathology)
31. (cytopath? OR histopath? OR immunohisto? OR histocyto? OR immunocyto?)
32. #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31
33. #16 AND #23 AND #32
34. 'limit 33 to' ('english language' AND 'yr="2015 -Current"')
35. 'limit 34 to' (comment OR editorial OR letter OR news OR 'newspaper article' OR 'observational study, veterinary' OR 'personal narrative')
36. #34 NOT #35

Cochrane Search String:

#1 MeSH descriptor: [Diagnostic Errors] explode all trees
#2 (delayed diagnosis OR incidental findings OR overdiagnosis):ti,ab,kw
#3 MeSH descriptor: [Near Miss, Healthcare]
#4 ("incidental findings"):ti,ab,kw
#5 ((over?diagnos* or mis?diagnos*)):ti,ab,kw
#6 (error* NEAR/2 (reduc* OR omission* OR commission* OR communication* OR consultation* OR critical)):ti,ab,kw
#7 (diagnos* NEAR/1 (miss or missed or mis or over)):ti,ab,kw

#8 ("near event*"):ti,ab,kw
#9 ("near miss*"):ti,ab,kw
#10 (event* NEAR/1 (sentinal OR adverse OR critical)):ti,ab,kw
#11 ((report* NEAR/1 amend*)):ti,ab,kw
#12 ("second opinion*"):ti,ab,kw
#13 (diagno* NEAR/2 (error* OR delay* OR accuracy OR uncertain*)):ti,ab.kw
#14 ((intra?operative or inter?institutional) NEAR/1 (consult* OR communication*)):ti,ab,kw
#15 ((observer OR interobserver) NEAR/1 (variation* OR variability OR agreements OR agreement OR consultat* or reproduc*)):ti,ab,kw
#16 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15
#17 MeSH descriptor: [Reproducibility of Results] explode all trees
#18 (quality NEXT/1 (assurance or improv*)):ti,ab,kw
#19 ("discordant rate*"):ti,ab,kw
#20 ("miss rate*"):ti,ab,kw
#21 (error* NEXT/1 (rate* OR reduct* OR detect*)):ti,ab,kw
#22 ((imprecision* OR concordance OR discordance OR reproducibility OR discrepan* OR ambigu* OR indeterm*)):ti,ab,kw
#23 #17 OR #18 OR #19 OR #20 OR #21 OR #22
#24 MeSH descriptor: [Pathology, Surgical] this term only
#25 MeSH descriptor: [Histocytochemistry] explode all trees
#26 MeSH descriptor: [Cytodiagnosis] explode all trees
#27 MeSH descriptor: [Histocytological Preparation Techniques] explode all trees
#28 MeSH descriptor: [Pathology, Molecular] this term only
#29 MeSH descriptor: [Cytological Techniques] explode all trees
#30 MeSH descriptor: [Histological Techniques] explode all trees
#31 (molecular diagnostic techniques):ti,ab,kw (Word variations have been searched)
#32 ((anatomic* OR molecular OR histo* OR surgical) NEAR/1 pathology):ti,ab,kw
#33 ((cytopath* OR histopath* OR immunohisto* OR histocyto* OR immunocyto*)):ti,ab,kw
#34 #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33
#35 #16 AND #23 AND #34

References

1. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol*. 2016;75:40-46. doi:10.1016/j.jclinepi.2016.01.021

2. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc*. 2016;104(3):240-243. doi:10.3163/1536-5050.104.3.014

3. National Institute for Health and Care Research. PROSPERO international prospective register of systematic reviews. Accessed October 9, 2025. https://www.crd.york.ac.uk/prospero/

4. Shea BJ, Reeves BC, Wells G et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358:j4008. doi:10.1136/bmj.j4008

5. Sterne JA, Hernan MA, Reeves BC et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919. doi:10.1136/bmj.i4919

6. Whiting PF, Rutjes AW, Westwood ME et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536. doi:10.7326/0003-4819-155-8-201110180-00009

7. Alonso-Coello P, Schunemann HJ, Moberg J et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ*. 2016;353:i2016. doi:10.1136/bmj.i2016

8. Gerhard R, Boerner SL. Evaluation of indeterminate thyroid cytology by second-opinion diagnosis or repeat fine-needle aspiration: which is the best approach? *Acta Cytol*. 2015;59(1):43-50. doi:10.1159/000369332

9. Adhya AK, Kar M, Mohanty R. Touch Imprint Cytology: A rapid and accurate method for diagnosis of oral cancer. *Acta Cytol*. 2019;63(5):411-416. doi:10.1159/000500006

10. Goyal A, Abdul-Karim FW, Yang B, Patel JB, Brainard JA. Interobserver agreement in the cytologic grading of atypia in neoplastic pancreatic mucinous cysts with the 2-tiered approach. *Cancer Cytopathol*. 2016;124(12):909-916. doi:10.1002/cncy.21767

11. Groen R, Abe K, Yoon HS et al. Application of microscope-based scanning software (Panoptiq) for the interpretation of cervicovaginal cytology specimens. *Cancer Cytopathol*. 2017;125(12):918-925. doi:10.1002/cncy.21921

12. Gru AA, Kim J, Pulitzer M et al. The use of central pathology review with digital slide scanning in advanced-stage mycosis fungoides and sezary yndrome: a multi-institutional and international pathology study. *Am J Surg Pathol*. 2018;42(6):726-734. doi:10.1097/PAS.0000000000001041

13. Jimeno M, Domingo A, Salas I et al. Pathologist experience and concordance in the diagnosis of dysplasia in long-standing inflammatory bowel disease. *Am J Surg Pathol*. 2020;44(7):955-961. doi:10.1097/PAS.0000000000001475

14. Layfield LJ, Hammer RD, Frazier SR et al. Impact of consensus conference review on diagnostic disagreements in the evaluation of cervical biopsy specimens. *Am J Clin Pathol*. 2017;147(5):473-476. doi:10.1093/ajcp/aqx024

15. Magalhaes JC, Azara CZS, Tavares S, Manrique EJC, Amaral RG. Impact of implementing 100% rapid review as a quality control tool in cervical cytology. *Acta Cytol*. 2018;62(2):115-120. doi:10.1159/000487426

16. Agrawal N, Kothari K, Tummidi S, Sood P, Agnihotri M, Shah V. Fine-needle aspiration biopsy cytopathology of breast lesions using the International Academy of Cytology

Yokohama System and Rapid On-Site Evaluation: a single-institute Experience. *Acta Cytol*. 2021;65(6):463-477. doi:10.1159/000518375

17. Amin A, Carr N. Diagnostic concordance in cases of appendiceal mucinous neoplasia referred to a tertiary referral centre. *J Clin Pathol*. 2019;72(9):639-641. doi:10.1136/jclinpath-2019-205945

18. Bailey GE, Graham A, Kahler J et al. The value of second-opinion consultation in nongynecologic cytopathology. *Am J Clin Pathol*. 2022;157(5):724-730. doi:10.1093/ajcp/aqab182

19. Balasubramanian P, Chandrashekar L, Thappa DM et al. A retrospective audit of skin biopsies done in a tertiary care center in India. *Int J Dermatol*. 2015;54(8):939-943. doi:10.1111/ijd.12718

20. Banet N, Levinson KL, VandenBussche CJ. Concordance of expert consultation diagnoses in the review of pelvic washing specimens. *Acta Cytol*. 2015;59(6):452-6. doi:10.1159/000443715

21. Bellevicine C, Migliatico I, Vigliar E, Serra N, Troncone G. Intra-institutional second opinion diagnosis can reduce unnecessary surgery for indeterminate thyroid FNA: a preliminary report on 34 cases. *Cytopathology*. 2017;28(4):254-258. doi:10.1111/cyt.12431

22. Bhoyrul B, Brent G, Elliott F et al. Pathological review of primary cutaneous malignant melanoma by a specialist skin cancer multidisciplinary team improves patient care in the UK. *J Clin Pathol*. 2019;72(7):482-486. doi:10.1136/jclinpath-2019-205767

23. Choi KY, Amit M, Tam S et al. Clinical implication of diagnostic and histopathologic discrepancies in sinonasal malignancies. *Laryngoscope*. 2021;131(5):E1468-E1475. doi:10.1002/lary.29102

24. Confortini M, Di Stefano C, Biggeri A et al. Daily peer review of abnormal cervical smears in the assessment of individual practice as an additional method of internal quality control. *Cytopathology*. 2016;27(1):35-42. doi:10.1111/cyt.12195

25. Endo C, Nakashima R, Taguchi A et al. Inter-rater agreement of sputum cytology for lung cancer screening in Japan. *Diagn Cytopathol*. 2015;43(7):545-550. doi:10.1002/dc.23253

26. Galli G, Trama A, Abate-Daga L, Brambilla M, Garassino MC, Fabbri A. Accuracy of pathologic diagnosis for thymic epithelial tumors: a brief report from an Italian reference Center. *Lung Cancer*. 2020;146:66-69. doi:10.1016/j.lungcan.2020.05.007

27. Giunchi F, Jordahl K, Bollito E et al. Interpathologist concordance in the histological diagnosis of focal prostatic atrophy lesions, acute and chronic prostatitis, PIN, and prostate cancer. *Virchows Arch*. 2017;470(6):711-715. doi:10.1007/s00428-017-2123-1

28. Khazai L, Middleton LP, Goktepe N, Liu BT, Sahin AA. Breast pathology second review identifies clinically significant discrepancies in over 10% of patients. *J Surg Oncol*. 2015;111(2):192-197. doi:10.1002/jso.23788

29. Kuijpers CC, Burger G, Al-Janabi S, Willems SM, van Diest PJ, Jiwa M. Improved quality of patient care through routine second review of histopathology specimens prior to multidisciplinary meetings. *J Clin Pathol*. 2016;69(10):866-871. doi:10.1136/jclinpath-2015-203488

30. Layfield LJ, Schmidt RL, Chadwick BE, Esebua M, Witt BL. Interobserver reproducibility and agreement with original diagnosis in the categories "atypical" and "suspicious for malignancy" for bile and pancreatic duct brushings. *Diagn Cytopathol*. 2015;43(10):797-801. doi:10.1002/dc.23305

31. Lohman ME, Grekin RC, North JP, Neuhaus IM. Impact of second-opinion dermatopathology reviews on surgical management of malignant neoplasms. *J Am Acad Dermatol*. 2021;84(5):1385-1392. doi:10.1016/j.jaad.2020.12.022

32. Mullin MH, Brierley DJ, Speight PM. Second opinion reporting in head and neck pathology: the pattern of referrals and impact on final diagnosis. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2015;119(6):656-660. doi:10.1016/j.oooo.2014.12.020

33. Orlando L, Viale G, Bria E et al. Discordance in pathology report after central pathology review: implications for breast cancer adjuvant treatment. *Breast*. 2016;30:151-155. doi:10.1016/j.breast.2016.09.015

34. Patrawala S, Maley A, Greskovich C et al. Discordance of histopathologic parameters in cutaneous melanoma: clinical implications. *J Am Acad Dermatol*. 2016;74(1):75-80. doi:10.1016/j.jaad.2015.09.008

35. Peck M, Moffat D, Latham B, Badrick T. Review of diagnostic error in anatomical pathology and the role and value of second opinions in error prevention. *J Clin Pathol*. 2018;71(11):995-1000. doi:10.1136/jclinpath-2018-205226

36. Shubhasini AR, Praveen BN, Usha H, et al. Inter- and intra-observer variability in diagnosis of oral dysplasia. *Asian Pac J Cancer Prev*. 2017;18(12):3251-3254. doi:10.22034/APJCP.2017.18.12.3251

37. Ronchi A, Pagliuca F, Zito Marino F et al. Second diagnostic opinion by experienced dermatopathologists in the setting of a referral regional melanoma unit significantly improves the clinical management of patients with cutaneous melanoma. *Front Med (Lausanne)*. 2020;7:568946. doi:10.3389/fmed.2020.568946

38. Soofi Y, Khoury T. Inter-institutional pathology consultation: the importance of breast pathology subspecialization in a setting of tertiary cancer center. *Breast J*. 2015;21(4):337-344. doi:10.1111/tbj.12420

39. Stewart CJR, Stewart LM, Holman CDJ et al. Value of pathology review in a population-based series of ovarian tumors. *Int J Gynecol Pathol*. 2017;36(4):377-385. doi:10.1097/PGP.0000000000000342

40. Vats K, Spafford M, Groot G et al. Moving towards the optimization of diagnosis for patients with sarcoma: a 10-year review of externally consulted sarcoma cases in a general anatomical pathology service. *Ann Diagn Pathol*. 2022;60:151958. doi:10.1016/j.anndiagpath.2022.151958

41. McHugh KE, Bird P, Sturgis CD. Concordance of breast fine needle aspiration cytology interpretation with subsequent surgical pathology: an 18-year review from a single sub-Saharan African institution. *Cytopathol*. 2019;30(5):519-525. doi:10.1111/cyt.12696

42. Middleton LP, Mayo TL, Spinks TE et al. The value of secondary pathology review. *J Clin Oncol*. 2016: 7-7. doi: 10.1200/jco.2016.34.7_suppl.7

43. Raju K, Sunanda SM, Rajanna VS. Cytology-histology correlation of gynecologic and non-gynecologic samples at a tertiary health care center: a two-year study. *Biomed Res and Ther*. 2019;6(4):3096-3106. 10.15419/bmrat.v6i4.534

44. Solivas J, Diwa MH. Rate of Diagnostic change in surgical pathology reports after mandatory intradepartmental peer review in a tertiary hospital in the Philippines: a retrospective study. *Acta Med Philipp*. 2024;58(16):42-49. doi:10.47895/amp.v58i16.3493

45. Boler AK, Roy S, Chakraborty A, Bandyopadhyay A. Reproducibility of the "International Academy of Cytology Yokohama System for Reporting Breast Cytology" - a retrospective analysis of 70 cases. *J Cytol*. 2022;39(4):159-162. doi:10.4103/joc.joc_203_21

46. Nagose VB, Kamal MM, Kathuria SS, Laddhad SA. The Indian Academy of Cytologists Guidelines for Reporting Serous Effusion (IACGRSE): an interobserver agreement analysis. *J Cytol*. 2022;39(3):121-125. doi:10.4103/joc.joc_50_22

47. Lezcano C, Berwick M, Luo L et al. Interobserver agreement in the histopathological classification of desmoplastic melanomas. *Pathology*. 2023;55(2):223-226. doi:10.1016/j.pathol.2022.12.347

48. Richards R, Agarwal S. Atypical squamous verrucous lesions of the oral cavity: challenges in interpretation of small incisional biopsies. *Head Neck Pathol*. 2023;17(3):607-617. doi:10.1007/s12105-023-01558-6

49. Sripodok S, Benjakul N. Interobserver variability in inconclusive diagnostic categories of thyroid fine needle aspiration cytology: an urban-based tertiary hospital experience. *Ann Diagn Pathol*. 2023;63:152083. doi:10.1016/j.anndiagpath.2022.152083

50. Woeste MR, Jacob K, Duff MB et al. Impact of routine expert breast pathology consultation and factors predicting discordant diagnosis. *Surg Oncol*. 2022;45:101860. doi:10.1016/j.suronc.2022.101860

51. Rakheja G, Handa U, Punia RS, Attri AK. Fine-needle aspiration cytology in soft tissue tumors-5-year institutional experience. *Diagn Cytopathol*. 2022;50(10):463-470. doi:10.1002/dc.25017

52. Shinohara S, Bychkov A, Munkhdelger J et al. Substantial improvement of histopathological diagnosis by whole-slide image-based remote consultation. *Virchows Arch*. 2022;481(2):295-305. doi:10.1007/s00428-022-03327-2

53. Torous VF, Pitman MB, Heher YK. The value of monitoring amended reports in cytopathology quality programs: a biennial review. *Cancer Cytopathol*. 2022;130(11):860-871. doi:10.1002/cncy.22607

54. Gupta K, Maitra D, Gowrishankar S. Whole slide imaging vs eyeballing: the future in quantification of tubular atrophy in routine clinical practice. *Indian J Nephrol*. 2022;32(2):151-155. doi:10.4103/ijn.IJN_333_20

55. Wellauer H, Studer G, Bode-Lesniewska B, Fuchs B. Time and accuracy to establish the diagnosis of soft tissue tumors: a comparative analysis from the Swiss Sarcoma Network. *Sarcoma*. 2022;2022:7949549. doi:10.1155/2022/7949549

56. Elmore JG, Barnhill RL, Elder DE et al. Pathologists' diagnosis of invasive melanoma and melanocytic proliferations: observer accuracy and reproducibility study. *BMJ*. 2017;357:j2813. doi:10.1136/bmj.j2813

57. Piepkorn MW, Longton GM, Reisch LM et al. Assessment of second-opinion strategies for diagnoses of cutaneous melanocytic lesions. *JAMA Netw Open*. 2019;2(10):e1912597. doi:10.1001/jamanetworkopen.2019.12597

58. van der Wel MJ, Coleman HG, Bergman J et al. Histopathologist features predictive of diagnostic concordance at expert level among a large international sample of pathologists diagnosing barrett's dysplasia using digital pathology. *Gut*. 2020;69(5):811-822.doi:10.1136/gutjnl-2019-318985

59. Nethanel A, Kyprianou C, Barzilai A et al. The Implications of a dermatopathologist's report on melanoma diagnosis and treatment. *Life (Basel)*. 2023;13(9):24. doi:10.3390/life13091803

60. Kuijpers CC, Visser M, Sie-Go DM et al. Improved cytodiagnostics and quality of patient care through double reading of selected cases by an expert cytopathologist. *Virchows Arch*. 2015;466(6):617-624. doi:10.1007/s00428-015-1738-3

61. Abro S, Nomani L, Wojcik EM, Pambuccian SE, Chatt G, Barkan GA. Outcome analysis and negative predictive value of the "unsatisfactory/nondiagnostic" category of The Paris System for Reporting Urinary Cytology. *J Am Soc Cytopathol*. 2021;10(1):64-70. doi:10.1016/j.jasc.2020.09.014

62. Agrawal M, Uppin MS, Uppin SG, Challa S, Agrawal S, Dharmrakshak AK. Thymoma diagnosis and categorization in the current scenario: morphological analysis based on interobserver variability. *Ann Thorac Med*. 2020;15(2):90-94. doi:10.4103/atm.ATM_350_19

63. Ahmad Z, Idrees R, Uddin N, Ahmed A, Fatima S. Errors in surgical pathology reports: a study from a major center in Pakistan. *Asian Pac J Cancer Prev*. 2016;17(4):1869-1874. doi:10.7314/apjcp.2016.17.4.1869

64. Cocks M, Gru AA. Quality assurance in dermatopathology: a review of report amendments. *J Cutan Pathol*. 2021;48(1):34-40. doi:10.1111/cup.13827

65. El Sharouni MA, Laeijendecker AE, Suijkerbuijk KP et al. High discordance rate in assessing sentinel node positivity in cutaneous melanoma: expert review may reduce unjustified adjuvant treatment. *Eur J Cancer*. 2021;149:105-113. doi:10.1016/j.ejca.2021.03.001

66. Gordetsky J, Collingwood R, Lai WS, Del Carmen Rodriquez Pena M, Rais-Bahrami S. Second opinion expert pathology review in bladder cancer: implications for patient care. *Int J Surg Pathol*. 2018;26(1):12-17. doi:10.1177/1066896917730903

67. Hohnen H, Dessauvagie B, Hardie M, McCallum D, Oehmen R, Latham B. Diagnostic concordance among pathologists interpreting breast core biopsies on secondary review over a 1-year period at an Australian tertiary hospital. *Breast J*. 2021;27(8):664-670. doi:10.1111/tbj.14267

68. Johnson SM, Samulski TD, O'Connor SM et al. Clinical and financial implications of second-opinion surgical pathology review. *Am J Clin Pathol*. 2021;156(4):559-568. doi:10.1093/ajcp/aqaa263

69. Layfield LJ, Frazier SR. Quality assurance of anatomic pathology diagnoses: comparison of alternate approaches. *Pathol Res Pract*. 2017;213(2):126-129. doi:10.1016/j.prp.2016.11.007

70. Yue SYP, Lucas SB, Brown M, Chiodini PL, Walker SL, Mahadeva U. Utility of an infectious and tropical disease histopathology diagnostic review service. *J Clin Pathol*. 2020;73(12):836-839. doi:10.1136/jclinpath-2020-206546

71. Elmore JG, Longton GM, Carney PA et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA*. 2015;313(11):1122-1132. doi:10.1001/jama.2015.1405

72. Govind D, Jen KY, Matsukuma K et al. Improving the accuracy of gastrointestinal neuroendocrine tumor grading with deep learning. *Sci Rep*. 2020;10(1):11064. doi:10.1038/s41598-020-67880-z

73. VandenBussche CJ, Adams C, Ali SZ, Olson MT. Cytotechnologist performance for screening Hurthle Cell Atypia in indeterminate thyroid fine-needle aspirates. *Acta Cytol*. 2015;59(5):377-383. doi:10.1159/000441939

74. Witt BL, Cohen MB, Chadwick BE, Stephenson PD, Abasolo P, Schmidt RL. Cytomorphology is often insufficient to categorize non-small-cell lung carcinoma on FNA specimens. *Diagn Cytopathol*. 2016;44(2):73-79. doi:10.1002/dc.23387

75. Guyatt G, Oxman AD, Akl EA et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64(4):383-394. doi:10.1016/j.jclinepi.2010.04.026

76. Al-Adnani M, Marnerides A, George S, Nasir A, Weber MA. "Delayed Villous Maturation" in placental reporting: concordance among consultant pediatric pathologists at a single specialist center. *Pediatr Dev Pathol*. 2015;18(5):375-379. doi:10.2350/12-02-1604-OA.1

77. Azara CZ, Manrique EJ, Tavares SB, Alves de Souza NL, Magalhaes JC, Amaral RG. Reproducibility of cervical cytopathology following an intervention by an external quality control laboratory. *Diagn Cytopathol*. 2016;44(4):305-310. doi:10.1002/dc.23445

78. Boennelycke M, Peters EEM, Leon-Castillo A et al. Prognostic impact of histological review of high-grade endometrial carcinomas in a large Danish cohort. *Virchows Arch*. 2021;479(3):507-514. doi:10.1007/s00428-021-03133-2

79. Bulten W, Balkenhol M, Belinga JA et al. Artificial intelligence assistance significantly improves Gleason grading of prostate biopsies by pathologists. *Mod Pathol*. 2021;34(3):660-671. doi:10.1038/s41379-020-0640-y

80. Crescenzi A, Trimboli P, Basolo F et al. Exploring the inter-observer agreement among the members of the Italian Consensus for the Classification and Reporting of Thyroid Cytology. *Endocr Pathol*. 2020;31(3):301-306. doi:10.1007/s12022-020-09636-z

81. de Morais LSF, Magalhaes JC, Braga IDS, Marega LA, Tavares S, Amaral RG. Performance of laboratories after 10 years of participating in external quality monitoring in cervical cytology. *Acta Cytol*. 2020;64(3):224-231. doi:10.1159/000502433

82. Dessauvagie BF, Lee AHS, Meehan K et al. Interobserver variation in the diagnosis of fibroepithelial lesions of the breast: a multicentre audit by digital pathology. *J Clin Pathol*. 2018;71(8):672-679. doi:10.1136/jclinpath-2017-204977

83. Gavrielides MA, Miller M, Hagemann IS et al. Clinical decision support for ovarian carcinoma subtype classification: a pilot observer study with pathology trainees. *Arch Pathol Lab Med*. 2020;144(7):869-877. doi:10.5858/arpa.2019-0390-OA

84. Grevenkamp F, Kommoss F, Kommoss F et al. Second opinion expert pathology in endometrial cancer: potential clinical implications. *Int J Gynecol Cancer*. 2017;27(2):289-296. doi:10.1097/IGC.0000000000000870

85. Hekler A, Utikal JS, Enk AH et al. Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. *Eur J Cancer*. 2019;118:91-96. doi:10.1016/j.ejca.2019.06.012

86. Hernandez-Prera JC, Machado RA, Asa SL et al. Pathologic reporting of tall-cell variant of papillary thyroid cancer: have we reached a consensus? *Thyroid*. 2017;27(12):1498-1504. doi:10.1089/thy.2017.0280

87. Hescot S, Sheikh-Alard H, Kordahi M et al. Impact of expert review of histological diagnosis of papillary and follicular thyroid cancer. *Endocrine*. 2021;72(3):791-797. doi:10.1007/s12020-020-02531-x

88. Kang HJ, Kwon SY, Kim A et al. A multicenter study of interobserver variability in pathologic diagnosis of papillary breast lesions on core needle biopsy with WHO classification. *J Pathol Transl Med*. 2021;55(6):380-387. doi:10.4132/jptm.2021.07.29

89. Lobo C, Costa J, Petronilho S, Monteiro P, Leca L, Schmitt F. Cytohistological correlation in serous effusions using the newly proposed International System for Reporting Serous Fluid Cytopathology: experience of an oncological center. *Diagn Cytopathol*. 2021;49(5):596-605. doi:10.1002/dc.24440

90. Lucas M, Jansen I, Savci-Heijink CD et al. Deep learning for automatic Gleason pattern classification for grade group determination of prostate biopsies. *Virchows Arch*. 2019;475(1):77-83. doi:10.1007/s00428-019-02577-x

91. Makela K, Hodgson U, Piilonen A et al. Analysis of the Histologic Features associated with interobserver variation in idiopathic pulmonary fibrosis. *Am J Surg Pathol*. 2018;42(5):672-678. doi:10.1097/PAS.0000000000001031

92. Marshall C, Mounzer R, Hall M et al. Suboptimal agreement among cytopathologists in diagnosis of malignancy based on endoscopic ultrasound needle aspirates of solid pancreatic lesions: a validation study. *Clin Gastroenterol Hepatol*. 2018;16(7):1114-1122 e2. doi:10.1016/j.cgh.2017.09.013

93. Mastracci L, Piol N, Molinaro L et al. Interobserver reproducibility in pathologist interpretation of columnar-lined esophagus. *Virchows Arch*. 2016;468(2):159-67. doi:10.1007/s00428-015-1878-5

94. Matheus WE, Ferreira U, Brandao EA, Ferruccio AA, Billis A. The importance of histopathologic review of biopsies in patients with prostate cancer referred to a tertiary uro-oncology center. *Int Braz J Urol*. 2019;45(1):32-37. doi:10.1590/S1677-5538.IBJU.2018.0099

95. Matsuda A, Kawabata H, Tohyama K et al. Interobserver concordance of assessments of dysplasia and blast counts for the diagnosis of patients with cytopenia: from the Japanese central review study. *Leuk Res*. 2018;74:137-143. doi:10.1016/j.leukres.2018.06.003

96. Mazariegos GV, Shneider BL, Shemesh E et al. Approaches to research determination of late acute cellular rejection in pediatric liver transplant recipients. *Liver Transpl*. 2021;27(1):106-115. doi:10.1002/lt.25903

97. Packer MDC, Ravinsky E, Azordegan N. Patterns of error in interpretive pathology. *Am J Clin Pathol*. 2022;157(5):767-773. doi:10.1093/ajcp/aqab190

98. Rakha EA, Ahmed MA, Aleskandarany MA et al. Diagnostic concordance of breast pathologists: lessons from the National Health Service Breast Screening Programme Pathology External Quality Assurance Scheme. *Histopathology*. 2017;70(4):632-642. doi:10.1111/his.13117

99. Schollenberg EL, Sapp HL, Huang WY. Inaccurate and incomplete diagnoses of malignant polyps as a cause of pathologic tumor stage T0 colectomy. *Ann Diagn Pathol*. 2015;19(1):16-19. doi:10.1016/j.anndiagpath.2014.11.003

100. Strosberg C, Gibbs J, Braswell D et al. Second opinion reviews for cancer diagnoses in anatomic pathology: a comprehensive cancer center's experience. *Anticancer Res*. 2018;38(5):2989-2994. doi:10.21873/anticanres.12551

101. Szecsei CM, Oxley JD. Errors in prostate core biopsy diagnosis in an era of specialisation and double reporting. *J Clin Pathol*. 2021;74(5):327-330. doi:10.1136/jclinpath-2020-206726

102. van der Wel MJ, Klaver E, Pouw RE et al. Significant variation in histopathological assessment of endoscopic resections for Barrett's neoplasia suggests need for consensus reporting: propositions for improvement. *Dis Esophagus*. 2021;34(12):24. doi:10.1093/dote/doab034

103. van Santvoort BWH, van Leenders G, Kiemeney LA et al. Histopathological re-evaluations of biopsies in prostate cancer: a nationwide observational study. *Scand J Urol*. 2020;54(6):463-469. doi:10.1080/21681805.2020.1806354

104. van Seijen M, Jozwiak K, Pinder SE et al. Variability in grading of ductal carcinoma in situ among an international group of pathologists. *J Pathol Clin Res*. 2021;7(3):233-242. doi:10.1002/cjp2.201

105. Vennalaganti P, Kanakadandi V, Goldblum JR et al. Discordance among pathologists in the United States and Europe in diagnosis of low-grade dysplasia for patients with Barrett's Esophagus. *Gastroenterology*. 2017;152(3):564-570 e4. doi:10.1053/j.gastro.2016.10.041

106. Villa I, Mathieu MC, Bosq J et al. Daily biopsy diagnosis in surgical pathology: concordance between light microscopy and whole-slide imaging in real-life conditions. *Am J Clin Pathol*. 2018;149(4):344-351. doi:10.1093/ajcp/aqx161

107. Villanacci V, Lorenzi L, Donato F et al. Histopathological evaluation of duodenal biopsy in the PreventCD project. an observational interobserver agreement study. *APMIS*. 2018;126(3):208-214. doi:10.1111/apm.12812

108. Yoshida H, Kikuchi A, Tsuda H et al. Discrepancies in pathological diagnosis of endometrial stromal sarcoma: a multi-institutional retrospective study from the Japanese Clinical Oncology Group. *Hum Pathol*. 2022;124:24-35. doi:10.1016/j.humpath.2022.03.007

109. Henna N, Fakhar SA, Akhter N et al. Relationship of cytological with histopathological examination of palpable thyroid nodule. *Pakistan J of Med Health Sci*. 2017;11(4):1395-1398.