Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

What Does a Positive Stain Measure? From Fixed Tissue to Living Cell Biology: Lessons from the Stomach

Birupaksha Biswas[1]

drbiswasassted.medicalglory@gmail.com

Burdwan Medical College and Hospital, West Bengal University of Health Sciences

mini-review

Keywords: Immunohistochemistry, antigen retrieval, protein localisation, membrane biology, quantitative histochemistry, tissue heterogeneity

Posted Date: November 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11611770/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

What Does a Positive Stain Measure? From Fixed Tissue to Living Cell Biology: Lessons from the Stomach

Article type: Review

Running title: From Fixed Epitopes to Cell Function

Author

Birupaksha Biswas[1], MBBS, MD (Pathology)

Affiliation

[1] Consultant Pathologist and Independent Researcher, Kolkata, West Bengal, India

Previous institutional appointment: Former Senior Resident, Department of Pathology, Burdwan Medical College and Hospital, West Bengal University of Health Sciences, Burdwan, West Bengal, India.

Corresponding author

Dr Birupaksha Biswas, MBBS, MD (Pathology)

Consultant Pathologist and Independent Researcher

Kolkata, West Bengal, India

Email: drbiswasassted.medicalglory@gmail.com

ORCID iD: https://orcid.org/0009-0007-5701-9131

Manuscript information

Manuscript category: Review

Keywords: Immunohistochemistry; antigen retrieval; protein localisation; membrane biology; quantitative histochemistry; tissue heterogeneity

Tables: 4 original conceptual tables

Figures: 2 original conceptual figures

Statements and Declarations

Funding

No external funding was received for this work.

Competing interests

The author declares no competing interests.

Ethics approval and consent to participate

Not applicable. This review does not report new investigations involving human participants or animals.

Consent for publication

Not applicable.

Data availability

No new datasets were generated or analysed for this article.

Author contributions

B.B. developed the conceptual framework, interpreted the scientific literature and assumes responsibility for the intellectual content of the manuscript.

Acknowledgements

None.

Generative AI-assisted technologies

AI-assisted technologies were used to generate the original conceptual illustrations presented in Figures 1 and 2 and to assist with language refinement. The author critically reviewed the content and assumes full responsibility for its scientific accuracy, interpretation and final presentation.

Abstract

Immunohistochemistry preserves the spatial relationship between molecular signals and tissue architecture, but its biological interpretation is less straightforward than its visual appearance suggests. Antibody binding in a fixed section establishes neither the accessibility of an epitope in living tissue nor the functional activity of the recognised protein. Membrane localisation, nuclear retention, staining intensity and cellular distribution are distinct measurements, each supporting different conclusions. Gastric adenocarcinoma provides an informative example because several clinically validated assays interrogate these properties within the same specimen. This critical narrative review examines the effects of fixation and antigen retrieval, membrane organisation and protein trafficking, cellular attribution, quantitative signal generation and spatial sampling on the interpretation of immunohistochemical findings. Building on established methodological literature, it proposes a measurement-to-claim framework that identifies the evidentiary gap between a histochemical observation and the biological conclusion drawn from it. The framework is intended neither as a new diagnostic score nor as an alternative to validated clinical thresholds. Its proposed contribution is to match specific interpretative uncertainties to complementary methods capable of addressing them. Applications extend beyond gastric pathology to receptor biology, developmental histology, neuroscience and quantitative tissue imaging. Its practical value remains to be tested experimentally. The central argument is that a reproducible stain becomes biologically informative only when the conclusion attributed to it is proportionate to the measurement performed.

Keywords: Immunohistochemistry; antigen retrieval; protein localisation; membrane biology; quantitative histochemistry; tissue heterogeneity

Introduction

A positive immunohistochemical stain can be among the most persuasive observations in experimental biology. It associates a molecular target with an identifiable cell, a subcellular compartment and an organised tissue. Nevertheless, the apparent precision of this localisation can conceal uncertainty about what has actually been measured. An antibody may detect a protein in a chemically fixed section without establishing its accessibility in a living cell. Membrane staining may identify a receptor without demonstrating its activation. Nuclear retention may indicate the presence of a repair protein without measuring its biochemical competence. These distinctions are familiar individually, but they are not always preserved when histochemical observations are translated into functional or clinical claims (Fritschy 2008; Shi et al. 2011).

This does not mean that immunohistochemistry (IHC) is inherently unreliable. Its value rests on the ability to associate molecular detection with tissue structure, often using assays whose analytical and clinical performance have been extensively validated. The difficulty arises when a conclusion extends beyond the property measured. Analytical validity, molecular localisation, functional inference and clinical utility are connected, but they are not interchangeable. A clinically useful threshold may be established through treatment-outcome evidence without representing a direct physical measurement of receptor activity or antibody access. Conversely, a technically reproducible signal may remain biologically ambiguous when its cellular origin, subcellular distribution or quantitative response is uncertain (Goldsmith et al. 2024).

Gastric adenocarcinoma provides a useful model for examining this problem. The same biopsy may be assessed for human epidermal growth factor receptor 2 (HER2), programmed death-ligand 1 (PD-L1), claudin-18.2 and mismatch-repair (MMR) proteins. These assays interrogate different biological properties and employ different interpretative conventions. Previous reviews have described gastric biomarker heterogeneity, diagnostic applications, digital methods and quality control; recent work has also examined claudin-18.2 accessibility beyond protein expression (Lee 2025; Oprea 2026; Deng et al. 2026). The present review does not claim novelty for those established limitations. Its contribution is a comparative framework that specifies what a measurement supports, identifies where inference exceeds observation and defines the independent evidence required to investigate the additional claim.

This critical narrative synthesis draws on selected foundational methodological studies, primary experimental investigations, professional guidance and clinically relevant biomarker research identified through targeted searches of PubMed-indexed records and publisher sources available up to 3 November 2026. The selection favours studies that clarify a particular measurement problem rather than those offering exhaustive coverage of gastric cancer. No protocol was registered, and no systematic eligibility screening, formal risk-of-bias assessment or quantitative pooling was undertaken. The conclusions are therefore interpretative and methodological; they establish neither the prevalence of diagnostic errors nor the comparative effectiveness of the proposed framework.

What fixation preserves—and what it changes

Immunohistochemistry begins with alterations to the material it seeks to examine. Fixation restricts degradation and preserves much of the microscopic organisation needed for tissue interpretation, but it also changes the chemical environment of proteins. Formaldehyde-mediated reactions may modify protein conformation, mask antibody-binding regions or affect the accessibility of an epitope. The resulting signal depends on antigen abundance, antibody specificity, fixation conditions, tissue processing, retrieval procedures and detection chemistry. A positive reaction is therefore an observation made under defined experimental conditions, rather than an unqualified representation of the protein's native state (Fritschy 2008; Shi et al. 2011).

Antigen retrieval illustrates this distinction particularly clearly. Heating can restore or improve the recognition of epitopes obscured during fixation and has substantially expanded the applicability of IHC. Yet improved staining after retrieval does not demonstrate that the original tissue contained more protein. It establishes that the antigen became more detectable under the revised conditions. Similarly, negative staining does not necessarily establish biological absence, particularly where antigen preservation, epitope recognition or antibody performance is uncertain. Adequate interpretation therefore requires controls capable of distinguishing analytical failure from genuine loss of detectable antigen, alongside evidence that the antibody recognises the intended target (Fritschy 2008; Goldsmith et al. 2024).

This distinction extends beyond diagnostic pathology. An investigator comparing neuronal receptor distribution, epithelial differentiation or nuclear transcription-factor expression may encounter differences produced partly by tissue preservation rather than underlying biology. Quantitative comparisons are particularly vulnerable when specimens have undergone different processing conditions. Standardisation can reduce such variability, but standardisation alone cannot establish that two antibodies recognise their respective targets with equivalent efficiency. Nor can it guarantee that an epitope exposed after processing was equally accessible in the living tissue.

The first evidentiary boundary therefore separates antigen detectability from the presence, abundance or accessibility of the native protein. Detection may provide strong evidence of protein localisation when the assay is appropriately controlled, but it cannot independently resolve questions of molecular conformation or function. The measurement must be interpreted within the experimental conditions in which it was obtained.

Where a protein is found is not necessarily what it does

A protein's location is biologically important, but localisation and activity are different measurements. This distinction is especially relevant to membrane proteins, whose properties depend on polarity, molecular

interactions, ligand access and intracellular trafficking. Epithelial membranes are not homogeneous surfaces. Apical, lateral and basal domains have different relationships with neighbouring cells, extracellular structures and circulating molecules. A membrane-associated signal therefore acquires meaning partly through the anatomical compartment in which it appears.

Claudin-18.2 demonstrates the importance of this organisation. The protein contributes to specialised epithelial junctions and is expressed in normal gastric mucosa as well as a subset of gastric carcinomas. Disruption of epithelial architecture during malignant transformation may alter the exposure of extracellular regions that are relatively shielded within intact junctional structures. Conventional IHC identifies detectable antigen in a processed section, but this measurement is not equivalent to determining whether a circulating therapeutic antibody can reach the target under physiological conditions. Antigen expression, membrane localisation, extracellular accessibility and therapeutic engagement are related properties rather than interchangeable observations (Nakayama et al. 2024; Deng et al. 2026).

This distinction must nevertheless be reconciled with established clinical evidence. Claudin-18.2 IHC has a validated role in treatment selection within specified clinical settings. The SPOTLIGHT trial employed a prespecified threshold of moderate-to-strong membranous staining in at least 75% of tumour cells and demonstrated benefit from zolbetuximab-containing treatment in the selected population (Shitara et al. 2023). That threshold is meaningful because it was used in a treatment setting supported by outcome data. It should not be interpreted as a direct measurement of the fraction of extracellular epitopes accessible to therapeutic antibodies, but neither should incomplete mechanistic understanding be used to undermine the clinical value of the validated assay (Fassan et al. 2024).

HER2 provides a related example in which membrane geometry, receptor abundance and biological activity must be distinguished. Gastric carcinoma commonly exhibits lateral or basolateral HER2 staining, reflecting epithelial architecture and necessitating organ-specific interpretative criteria (Bartley et al. 2016). A qualifying membrane pattern establishes detectable receptor expression within the relevant scoring system. It does not directly measure receptor phosphorylation, downstream signalling or the rates of receptor internalisation and recycling. These dynamic properties require methods designed to interrogate them. Experimental studies of antibody-induced HER2 trafficking demonstrate that receptor distribution can change following molecular engagement, underscoring why a fixed-tissue image cannot itself establish trafficking kinetics (Cheng et al. 2020).

Protein expression is also distinct from the molecular mechanism responsible for that expression. Nishida et al. (2015) used a simultaneous HER2 gene–protein assay to demonstrate intratumoural differences in receptor overexpression and gene amplification. Such approaches provide complementary information because gene copy number and protein localisation are separate biological measurements. The October 2026 update to the College of American Pathologists' gastro-oesophageal HER2 guidance continues to favour IHC with reflex in situ hybridisation for equivocal findings, while recognising the importance of specimen quality and heterogeneity (Bartley et al. 2026). Not every observation requires another test. Complementary testing is useful when it addresses an unresolved question relevant to the intended decision.

A similar boundary applies to intracellular proteins. Nuclear staining identifies detectable protein within the nucleus but does not necessarily demonstrate transcriptional activity, enzymatic competence or downstream molecular effects. MMR immunohistochemistry illustrates the distinction. Retention or loss of MLH1, PMS2, MSH2 and MSH6 in tumour-cell nuclei supports established assessment of MMR status when internal controls and validated interpretative criteria are satisfied. Nevertheless, retained immunoreactivity does not invariably guarantee preserved repair function, and apparent loss may occasionally reflect analytical conditions. Molecular microsatellite instability testing can resolve selected discrepant cases because it interrogates a related genomic phenotype rather than simply repeating the protein-expression measurement (Wang et al. 2022; Park et al. 2023).

These examples show why a positive stain should not automatically be converted into a functional claim. A receptor may be detected without its activity being measured; a junctional protein may be localised without its extracellular accessibility being quantified; and a nuclear protein may be retained without its biochemical

performance being directly demonstrated. Table 1 compares the properties measured by these assays. The differences do not diminish the value of clinically validated assays. They clarify the evidentiary basis of the conclusions those assays permit.

A positive signal must belong to the right cell

Tissues contain interacting cell populations whose molecular contributions cannot be inferred from aggregate positivity alone. Malignant epithelial cells, immune cells, vascular elements, fibroblasts and residual non-neoplastic structures may occupy the same microscopic field. Correct interpretation therefore requires an account of which cells contribute to the signal and whether they are included within the relevant scoring method.

PD-L1 assessment in gastric carcinoma demonstrates the importance of cellular attribution. The combined positive score (CPS) incorporates qualifying PD-L1-positive tumour cells, lymphocytes and macrophages, using viable tumour cells as the denominator. It is not simply the proportion of malignant epithelial cells expressing PD-L1. Two specimens may receive comparable scores despite markedly different distributions of tumour-cell and immune-cell staining. Such results can be clinically equivalent under a validated scoring system without describing identical cellular biology (Angerilli et al. 2023; Ahn et al. 2024).

The distinction has consequences for both interpretation and research. Positive macrophages may be difficult to separate from malignant cells in poorly differentiated tissue, while inflammation and tissue architecture can alter the distribution of qualifying cells. Assay-specific criteria and morphological correlation remain essential. At the same time, separating the cellular contributions to a composite score may permit investigators to ask whether particular tumour–immune configurations are associated with different biological responses. That question is legitimate, but the additional measurement must demonstrate value beyond the validated score before it can influence clinical selection. Differences between PD-L1 assays and scoring approaches reinforce the need to preserve the distinction between a clinically defined measurement and an experimental interpretation of its cellular components (Klempner et al. 2024).

Cellular attribution is equally important in nuclear and multiplex staining. Retained MMR expression in stromal or inflammatory cells serves as an internal control but cannot substitute for tumour-cell nuclear staining. In fluorescence microscopy, apparent overlap between two labelled proteins supports spatial proximity within the limits of optical resolution; it does not establish direct molecular interaction. Colocalisation depends on image resolution, signal-to-noise characteristics, background correction and the quantitative method used to assess spatial correspondence (Dunn et al. 2011). Thus, increasingly sophisticated imaging may improve localisation without resolving whether the detected molecules interact functionally.

The interpretative problem is not merely whether staining is present, but whether the signal has been assigned to the appropriate cell, compartment and spatial relationship. Failure at this stage can yield a highly reproducible measurement of an incorrectly defined biological object. Fig. 1 illustrates how broadly comparable staining impressions may arise from different cellular sources and spatial distributions, without implying equivalent measured staining areas or interchangeable assay-specific scores.

When a stain becomes a number

Digital imaging has improved the reproducibility and granularity of histochemical analysis. Software can quantify positive areas, segment cells, measure optical density and describe spatial relationships across whole-slide images. These capabilities strengthen tissue-based research, but they do not automatically make chromogenic staining a direct quantitative assay of protein concentration. Signal intensity depends on antigen recognition, amplification chemistry, chromogen deposition, section thickness, image acquisition and analytical settings. Non-linearity and saturation may prevent substantial differences in antigen abundance from producing proportionate differences in measured intensity (Kashyap et al. 2019).

Kashyap et al. (2019) addressed this problem through quantitative microimmunohistochemistry, using staining kinetics to obtain information unavailable from a conventional end-point image. The study demonstrates why quantitative interpretation requires knowledge of the assay's response behaviour. A numerical value is

informative only in relation to the conditions under which it was generated and the range over which it can be interpreted.

Instrument variation introduces another source of uncertainty. Rojansky et al. (2022) combined digital image analysis and quantitative bead standards to investigate inconsistent immunostaining and identified problems involving laboratory equipment, including variable electrical power delivery. Van Kempen et al. (2025) subsequently examined artificial intelligence-assisted monitoring of standardised cell-line controls across automated staining platforms and documented variation involving instruments and slide positions, particularly at lower expression levels. These findings demonstrate that weak or intermediate staining cannot always be attributed to biological differences or observer judgement. Analytical variation may precede interpretation.

Computer-assisted analysis can help identify such variation, but its own performance depends on the validity of the features selected for measurement. Image-analysis platforms such as QuPath facilitate reproducible segmentation and quantification, yet stable numerical outputs do not guarantee correct cellular attribution or clinically meaningful classification (Bankhead et al. 2017). Analytical precision and biological validity must therefore be evaluated separately.

Interobserver statistics offer a complementary example. In a multicentre assessment of 460 gastric and gastro-oesophageal junction cancer specimens, Zhu et al. (2026) reported an overall agreement proportion of 0.14 and Fleiss' κ of 0.41 for four-category HER2 IHC scoring by 15 pathologists. Reclassification into three categories increased the overall agreement proportion to 0.38 and κ to 0.50. These statistics describe different aspects of agreement and are not interchangeable measures of diagnostic accuracy. Nor does disagreement across all four categories imply equivalent unreliability at each clinically relevant decision boundary. Reliability depends on the classification question being asked.

Quantification therefore requires three safeguards: adequate control of the analytical response, correct identification of the measured object and interpretation appropriate to the intended use. Table 2 distinguishes these sources of uncertainty. Greater digital precision is valuable when it improves one or more of these properties; it is not evidence, by itself, that the underlying biological conclusion has become more reliable.

A section is not the whole tissue

An immunohistochemical result may be technically correct and appropriately interpreted yet remain unrepresentative of the biological system from which the specimen was obtained. Cells within the same tissue can differ in differentiation, protein expression, local microenvironment and molecular state. A section consequently describes a selected spatial compartment rather than an entire organ or disease process.

The problem is conspicuous in gastric carcinoma. HER2 amplification and protein expression may vary between tumour regions, producing discordance between limited samples and larger specimens (Nishida et al. 2015; Lee 2025). Claudin-18.2 expression may also differ between primary tumours and metastases. In a systematic review and meta-analysis of 896 paired samples, Dahlgaard et al. (2026) reported a pooled primary–metastatic discordance proportion of 21% (95% confidence interval 18–24%). The estimate concerns classification differences between sampled paired lesions under the included studies' methods and thresholds. It must not be transformed into a universal patient-level error rate.

Spatial heterogeneity should not be confused with technical variation. Two regions can yield different results because their biology differs, whereas repeated processing of the same material may produce variation through fixation, retrieval, staining or observer interpretation. The appropriate response depends on the source of uncertainty. Additional sampling may be informative when tissue representation is inadequate, but it cannot correct a systematically unstable assay. Conversely, improved calibration of the original section cannot demonstrate whether an unsampled lesion has a different molecular phenotype.

The clinical significance of spatial variation also depends on the decision being made. The 2026 HER2 guidance addresses specimen selection and circumstances in which repeat assessment may be appropriate, rather than requiring unrestricted multisite testing (Bartley et al. 2026). A research programme may reasonably investigate

spatial heterogeneity more extensively, but evidence of regional variation alone does not establish that additional sampling will improve patient outcomes.

This distinction extends across histochemistry. The distribution of a receptor in one epithelial compartment does not necessarily characterise the entire gland; a neuronal marker observed in one anatomical region cannot automatically be generalised throughout the nervous system; and a nuclear protein detected in one cell population may not represent its expression in neighbouring populations. In each case, the spatial unit being measured must be distinguished from the biological unit about which a conclusion is proposed.

Matching the claim to the measurement

The preceding evidence supports a framework that distinguishes four boundaries between a histochemical observation and its interpretation. The framework is not intended to replace established laboratory validation, diagnostic scoring or clinical decision-making. Its purpose is narrower: to identify the precise property that remains unmeasured when a conclusion extends beyond the original observation.

The first boundary concerns detectability. An antibody-dependent signal establishes recognition under defined experimental conditions, but not necessarily native molecular abundance or epitope accessibility. The second concerns topology. Localisation within a cellular compartment does not directly establish receptor activation, molecular interaction or dynamic trafficking. The third concerns representation. A result from particular cells or tissue regions may not describe the complete biological system. The fourth concerns inference. An analytically valid biomarker may support a clinically validated association without proving the mechanism responsible for that association.

These boundaries are distinct, although they may coexist. Improved antigen retrieval can address insufficient detectability but does not measure receptor activity. Multiplex imaging can clarify cellular attribution without demonstrating functional interaction. Additional sampling can improve spatial representation without correcting staining variability. Molecular testing may establish a genomic alteration but not necessarily the localisation of its protein product. The contribution of the framework is to make these differences explicit and to require that any proposed supplementary investigation address the particular uncertainty identified.

The approach rests on a straightforward methodological distinction. An observation reports what was measured. An inference connects that observation to another property through established evidence. A hypothesis proposes a relationship that remains to be tested. These categories should not be treated as interchangeable merely because they concern the same protein. Fig. 2 maps each of the four independent evidentiary boundaries to an appropriate class of complementary evidence; it is not a mandatory sequence of investigations or an alternative clinical algorithm.

The framework also guards against unnecessary investigation. A clinically validated assay can remain appropriate even when the molecular mechanism linking its result to treatment response is incompletely understood. Additional testing is justified only when it addresses a relevant unresolved question and possesses adequate analytical performance. The distinction between the existing measurement, the proposed claim and the possible complementary method is set out in Table 3.

The proposed contribution is not a new description of antigen masking, membrane polarity, staining heterogeneity or signal saturation. Each is established. It is the use of a common, claim-specific interpretative structure across assays that measure fundamentally different biological properties. Whether this structure improves scientific interpretation beyond existing methodological practice remains an empirical question.

What would establish the framework's value?

A conceptual framework should generate predictions capable of being tested independently. The proposal predicts that explicitly distinguishing the measured property from the intended biological conclusion will improve recognition of unsupported inferences and selection of appropriate complementary methods. This advantage has not been demonstrated and should not be presumed.

An initial evaluation could use standardised tissue and cell-line preparations assessed across laboratories under controlled fixation, retrieval and staining conditions. Conventional interpretation could be compared with an approach requiring explicit identification of the relevant measurement boundary. Independent reference methods would then determine whether differences in staining arose from altered antigen detectability, analytical variation or genuine differences in biological material. The primary endpoint should be correct attribution of the uncertainty, not merely reduced variation in staining intensity.

A second investigation could examine the distinction between localisation and accessibility. Standard fixed-tissue IHC for a selected membrane protein could be compared with validated measurements on non-permeabilised living cells, supported where appropriate by surface-binding or time-resolved trafficking studies. In claudin-18.2 research, the question would be whether an independent accessibility measurement explains biological or treatment-related variation beyond conventional expression and distribution. Demonstrating a difference between assays would be insufficient; the additional measurement would need to show that it captures a relevant property more accurately or provides meaningful incremental information.

A third approach could test cellular attribution and spatial representation. Blinded observers or validated image-analysis systems could assess the cellular contributors to a composite stain, while matched tissue regions would permit estimation of spatial variation separately from technical and interpretative variation. For PD-L1, a compartment-specific research analysis could be compared against established CPS assessment without displacing its validated clinical use. For heterogeneous membrane markers, additional sampling could be evaluated against prespecified classification or management-related endpoints rather than the mere detection of more heterogeneity.

Finally, a direct test of the interpretative framework could compare conventional expert review with review incorporating explicit measurement-to-claim assessment. Investigators would receive identical staining data and independently adjudicated reference information. Outcomes could include the frequency of unsupported mechanistic conclusions, agreement about the source of uncertainty and the appropriateness of proposed complementary methods. Any advantage would need to persist across observers, laboratories and biological contexts. Failure to improve interpretation, or greater complexity without clearer conclusions, would weaken any claim of practical utility. Table 4 summarises these proposed tests and their principal limitations.

Such experiments would also identify settings in which the framework is unnecessary. A routine specimen adequately assessed by a validated diagnostic algorithm may require no further investigation. The framework is most plausibly useful when the intended conclusion concerns protein behaviour, uncertain cellular attribution, tissue-wide representation or the interpretation of a quantitatively ambiguous signal.

Implications for histochemistry and cell biology

The measurement-to-claim problem extends beyond cancer biomarkers. In developmental biology, protein localisation is used to investigate differentiation and tissue organisation. In neuroscience, immunostaining identifies cellular populations and molecular compartments. In epithelial and renal biology, the distribution of junctional proteins and transporters helps define structural polarity. These applications depend on the same experimental distinction: localisation in a prepared specimen is an observation, whereas functional behaviour in the living system is an inference requiring suitable evidence.

Advances in multiplex imaging, automated segmentation and spatial analysis make this distinction increasingly important. These methods permit more detailed descriptions of tissue organisation, but additional measurements do not automatically establish molecular interactions or causal relationships. High-resolution spatial correspondence remains different from direct binding; relative signal intensity remains different from absolute molecular abundance; and a cross-sectional image remains different from a time-resolved observation.

The central requirement is not that every histochemical study should incorporate every available complementary method. It is that the method selected should be capable of answering the question being asked. Histochemistry is most informative when its spatial observations are combined, where necessary, with biochemical, molecular or dynamic experiments directed at the unresolved biological property.

Limitations and conclusions

This review is a selective conceptual synthesis; the proposed framework has not undergone experimental validation. It cannot establish that explicit measurement-to-claim assessment improves diagnostic accuracy, interobserver agreement, tissue utilisation or clinical outcomes. The studies discussed differ in specimen type, analytical method, biological target and intended use, preventing the derivation of a common quantitative estimate of interpretative error. Moreover, the underlying distinctions between antigen detection, localisation and function are established principles of experimental biology. The contribution proposed here is their comparative organisation into a framework for evaluating the evidentiary sufficiency of specific claims, not their discovery.

Immunohistochemistry measures antibody-dependent signals within prepared tissues. Its biological strength lies in relating those signals to cellular identity and spatial organisation. Its limitations arise when detectability is mistaken for native accessibility, localisation for function, numerical precision for molecular accuracy or a sampled region for the entire biological system. Clinically validated scoring systems remain authoritative within their intended settings, irrespective of whether every aspect of the underlying mechanism has been resolved.

A positive stain does not, by itself, explain cellular behaviour. It is a molecular observation situated within tissue architecture. Its scientific value depends on how closely the resulting claim reflects that observation and whether sufficient evidence supports any further inference.

References

Ahn S, Kwak Y, Kwon GY et al (2024) Interpretation of PD-L1 expression in gastric cancer: summary of a consensus meeting of Korean gastrointestinal pathologists. J Pathol Transl Med 58:103–116. https://doi.org/10.4132/jptm.2024.03.15

Angerilli V, Fassan M, Parente P et al (2023) A practical approach for PD-L1 evaluation in gastroesophageal cancer. Pathologica 115:57–70. https://doi.org/10.32074/1591-951X-836

Bankhead P, Loughrey MB, Fernández JA et al (2017) QuPath: open source software for digital pathology image analysis. Sci Rep 7:16878. https://doi.org/10.1038/s41598-017-17204-5

Bartley AN, Washington MK, Ventura CB et al (2016) HER2 testing and clinical decision making in gastroesophageal adenocarcinoma: guideline from the College of American Pathologists, American Society for Clinical Pathology, and the American Society of Clinical Oncology. Arch Pathol Lab Med 140:1345–1363. https://doi.org/10.5858/arpa.2016-0331-CP

Bartley AN, Ajani JA, Bridge JA et al (2026) Human epidermal growth factor receptor 2 testing and clinical decision-making in gastroesophageal adenocarcinoma: guideline update from the College of American Pathologists in collaboration with the American Society for Clinical Pathology. Arch Pathol Lab Med. Published online 21 October 2026. https://doi.org/10.5858/arpa.2026-0019-CP

Cheng J, Liang M, Carvalho MF et al (2020) Molecular mechanism of HER2 rapid internalization and redirected trafficking induced by anti-HER2 biparatopic antibody. Antibodies (Basel) 9:49. https://doi.org/10.3390/antib9030049

Dahlgaard NB, Tong Y, Shitara K, Fassan M, Mau-Sørensen M, Egebjerg K (2026) Claudin 18.2 expression—concordance between primary and metastatic lesions in gastric and gastroesophageal adenocarcinomas: a systematic review and meta-analysis. ESMO Gastrointest Oncol 13:100369. https://doi.org/10.1016/j.esmogo.2026.100369

Deng Y, Han J, Zhang X et al (2026) Beyond expression: molecular imaging of Claudin18.2 to decipher target accessibility and guide precision oncology. Pharmacol Res 231:108387. https://doi.org/10.1016/j.phrs.2026.108387

Dunn KW, Kamocka MM, McDonald JH (2011) A practical guide to evaluating colocalization in biological microscopy. Am J Physiol Cell Physiol 300–C742. https://doi.org/10.1152/ajpcell.00462.2010

Fassan M, Kuwata T, Matkowskyj KA, Röcken C, Rüschoff J (2024) Claudin-18.2 Immunohistochemical Evaluation in Gastric and Gastroesophageal Junction Adenocarcinomas to Direct Targeted Therapy: A Practical Approach. Mod Pathol 37:100589. https://doi.org/10.1016/j.modpat.2024.100589

Fritschy JM (2008) Is my antibody-staining specific? How to deal with pitfalls of immunohistochemistry. Eur J Neurosci 28:2365–2370. https://doi.org/10.1111/j.1460-9568.2008.06552.x

Goldsmith JD, Troxell ML, Roy-Chowdhuri S et al (2024) Principles of analytic validation of immunohistochemical assays: guideline update. Arch Pathol Lab Med 148–e153. https://doi.org/10.5858/arpa.2023-0483-CP

Kashyap A, Fomitcheva Khartchenko A, Pati P et al (2019) Quantitative microimmunohistochemistry for the grading of immunostains on tumour tissues. Nat Biomed Eng 3:478–490. https://doi.org/10.1038/s41551-019-0386-3

Klempner SJ, Cowden ES, Cytryn SL et al (2024) PD-L1 immunohistochemistry in gastric cancer: comparison of combined positive score and tumor area positivity across 28-8, 22C3, and SP263 assays. JCO Precis Oncol 8. https://doi.org/10.1200/PO.24.00230

Lee HS (2025) Spatial and temporal tumor heterogeneity in gastric cancer: discordance of predictive biomarkers. J Gastric Cancer 25:192–209. https://doi.org/10.5230/jgc.2025.25.e3

Nakayama I, Qi C, Chen Y, Nakamura Y, Shen L, Shitara K (2024) Claudin 18.2 as a novel therapeutic target. Nat Rev Clin Oncol 21:354–369. https://doi.org/10.1038/s41571-024-00874-2

Nishida Y, Kuwata T, Nitta H et al (2015) A novel gene-protein assay for evaluating HER2 status in gastric cancer: simultaneous analyses of HER2 protein overexpression and gene amplification reveal intratumoral heterogeneity. Gastric Cancer 18:458–466. https://doi.org/10.1007/s10120-014-0394-7

Oprea B (2026) Comprehensive overview of gastric cancer immunohistochemistry: key biomarkers, advanced detection methods, and perspectives. Medicina (Kaunas) 62:683. https://doi.org/10.3390/medicina62040683

Park Y, Nam SK, Seo SH et al (2023) Comprehensive study of microsatellite instability testing and its comparison with immunohistochemistry in gastric cancers. J Gastric Cancer 23:264–274. https://doi.org/10.5230/jgc.2023.23.e5

Rojansky R, Sompuram SR, Gomulia E, Natkunam Y, Troxell ML, Fernandez-Pol S (2022) Digital image analysis and quantitative bead standards in root cause analysis of immunohistochemical staining variability: a real-world example. Appl Immunohistochem Mol Morphol 30:477–485. https://doi.org/10.1097/PAI.0000000000001045

Shi SR, Shi Y, Taylor CR (2011) Antigen retrieval immunohistochemistry: review and future prospects in research and diagnosis over two decades. J Histochem Cytochem 59:13–32. https://doi.org/10.1369/jhc.2010.957191

Shitara K, Lordick F, Bang YJ et al (2023) Zolbetuximab plus mFOLFOX6 in patients with CLDN18.2-positive, HER2-negative, untreated, locally advanced unresectable or metastatic gastric or gastro-oesophageal junction adenocarcinoma (SPOTLIGHT): a multicentre, randomised, double-blind, phase 3 trial. Lancet 401:1655–1668. https://doi.org/10.1016/S0140-6736(23)00620-7

van Kempen S, Ghlowy Gerritsen WJ, Nguyen TQ et al (2025) Monitoring Immunohistochemical Staining Variations Using Artificial Intelligence on Standardized Controls. Lab Invest 105:104105. https://doi.org/10.1016/j.labinv.2025.104105

Wang C, Zhang L, Vakiani E, Shia J (2022) Detecting mismatch repair deficiency in solid neoplasms: immunohistochemistry, microsatellite instability, or both? Mod Pathol 35:1515–1528. https://doi.org/10.1038/s41379-022-01109-4

Zhu X, Yang L, Zheng Y et al (2026) Multi-institutional assessment of HER2 immunohistochemistry in gastric and esophagogastric junction cancer. iScience 29:115635. https://doi.org/10.1016/j.isci.2026.115635

Tables

Table 1. Biological properties measured by established immunohistochemical assays

Assay	Principal observation	Established interpretative use	Property not measured directly
HER2	Tumour-cell membrane receptor expression	Validated IHC/ISH-based classification	Receptor phosphorylation, trafficking or downstream signalling
PD-L1	Assay-defined staining of malignant and eligible immune cells	Clinically validated composite scoring	Independent functional contribution of each cellular population
Claudin-18.2	Qualifying tumour-cell membrane-associated antigen expression	Treatment-selection classification within validated settings	Extracellular epitope accessibility or therapeutic engagement in vivo
MMR proteins	Nuclear retention or loss of selected proteins	Assessment of MMR status	Direct biochemical DNA-repair activity

Notes: IHC, immunohistochemistry; ISH, in situ hybridisation; MMR, mismatch repair. Clinical eligibility continues to depend on the applicable validated assay and treatment indication.

Table 2. Sources of disagreement that should not be treated as equivalent

Source	Representative observation	Principal uncertainty
Epitope masking	Reduced staining after different fixation conditions	Detectability
Instrument variation	Different signal intensities across staining runs	Analytical reproducibility
Cellular attribution	Immune-cell staining contributing to a composite score	Identity of the measured cell population
Membrane organisation	Junctional antigen detected in a fixed section	Native accessibility
Observer disagreement	Different categories assigned to the same HER2 section	Classification reproducibility
Tissue heterogeneity	Different results in paired anatomical sites	Spatial representativeness
Molecular–protein discordance	Amplification without uniform protein localisation	Relationship between distinct molecular measurements

Notes: The categories may coexist and do not represent comparable rates of error. Their distinction determines which additional investigation, if any, is appropriate.

Table 3. Matching an unresolved biological claim to complementary evidence

Proposed claim	What conventional IHC provides	Potential complementary evidence	Remaining limitation
The epitope is accessible in living tissue	Detectability after processing	Validated live-cell or surface-binding experiments	Experimental accessibility may not establish clinical benefit
A receptor is functionally active	Protein localisation	Appropriate biochemical or dynamic functional assays	Results remain dependent on model and assay conditions
Two proteins interact	Spatial localisation or optical overlap	Validated interaction or proximity methods	Proximity does not invariably establish functional interaction
The result represents the whole tissue	Distribution in sampled material	Regionally appropriate sampling	Additional samples may remain incomplete
Digital intensity reflects abundance	Quantified image signal	Calibrated response measurements and controls	Absolute quantification requires suitable reference standards
A biomarker predicts benefit	Validated clinical classification	Incremental predictive studies	Association does not necessarily prove mechanism

Notes: These methods are examples for research or problem-solving, not mandatory supplementary clinical tests.

Table 4. Proposed experimental tests of the measurement-to-claim framework

Investigation	Principal endpoint	Necessary comparator	Finding that would weaken the framework's utility
Multicentre staining controls	Correct attribution of technical signal variation	Standard validation and quality assurance	No improvement in identifying sources of variation
Fixed-tissue versus live-cell accessibility	Incremental information about native antigen accessibility	Conventional IHC and independent binding reference	No relevant added biological information
Compartment-specific assessment	Additional information beyond aggregate scoring	Validated assay-specific score	No reproducible incremental value
Spatial sampling study	Improved representation or clinically important classification	Standard specimen selection	More detected variation without meaningful benefit
Blinded interpretative comparison	Reduction in unsupported biological claims	Standard expert interpretation	No improvement in interpretative accuracy or consistency

Notes: These experiments are proposed and have not been performed as part of this review. Any clinical benefit would require separate validation.

Figures

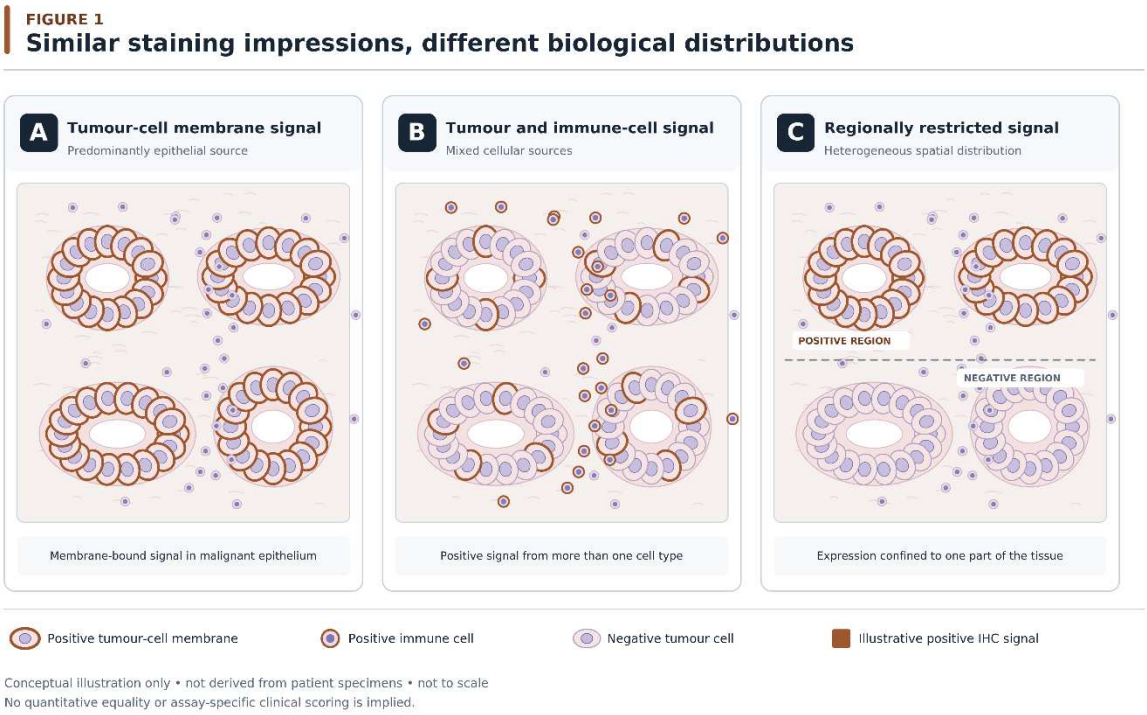


Figure 1. Similar immunohistochemical staining impressions, different biological distributions

Conceptual representations of predominantly tumour-cell membrane staining (A), staining contributed by both malignant epithelial and immune cells (B), and spatially restricted positivity within heterogeneous tissue (C). Brown denotes simulated positive immunohistochemical signal. The figure illustrates how cellular attribution and spatial distribution may differ despite broadly comparable staining impressions. The images are original schematic illustrations, not derived from patient specimens, and are not to scale. No quantitative equivalence or assay-specific clinical classification is implied.

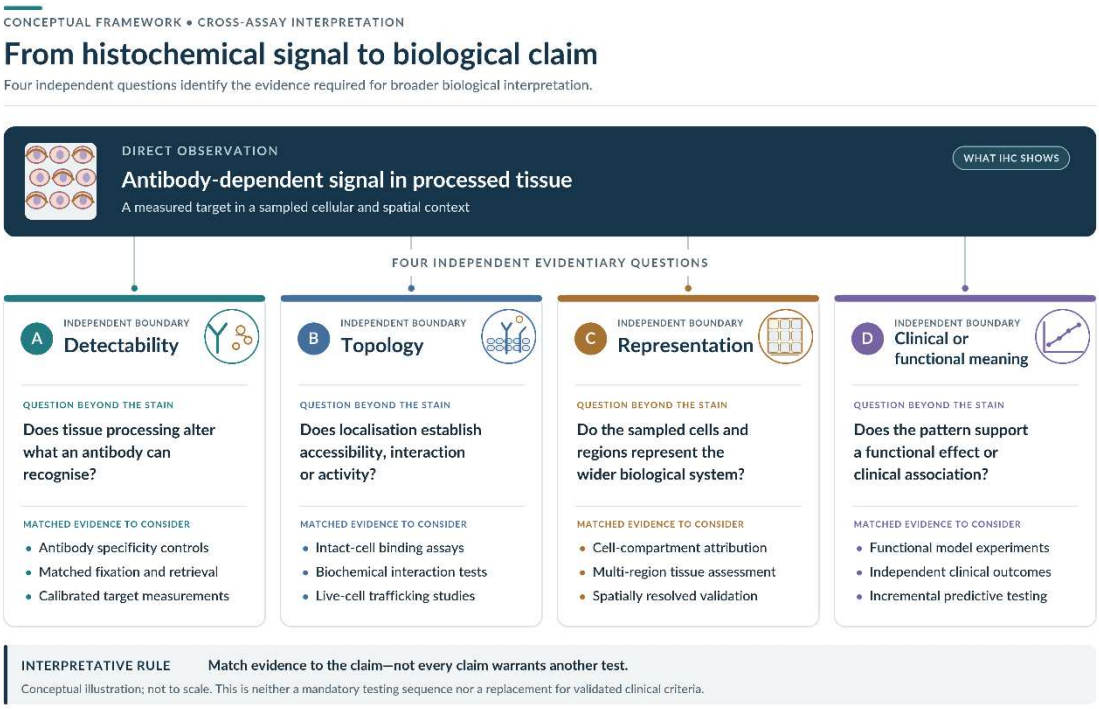


Figure 2. From histochemical measurement to biological claim: an evidence-matched framework

An observed immunohistochemical signal in processed tissue is linked to four independent evidentiary boundaries: epitope detectability (A), cellular and subcellular topology (B), spatial representativeness (C), and functional or clinical interpretation (D). Each boundary identifies a biological question that the original observation cannot resolve independently and illustrates complementary methods capable of investigating the additional claim. The framework is conceptual and does not prescribe a mandatory sequence of investigations, replace validated clinical scoring systems or establish a new treatment-eligibility algorithm. All illustrations are schematic and not to scale.