

Glial-Immune Microdomains in a Longitudinal Review

Abstract

This conceptual analysis examines spatiotemporal sampling in immune-state and tumor evidence synthesis, using tumor niches, immune exhaustion, treatment timing as the key comparison axes. The article develops a research agenda through critical comparison of published evidence. The focal evidence concerns why autoimmune activity can persist or recur when conventional inflammation suppression does not durably change underlying cellular and tissue states. A spatial snapshot is a sample of a changing process. Apparent niches can reflect sampling location, treatment timing, or transient cell states. Retain sampling coordinates and treatment chronology. Compare patterns across times and regions, with controls for preparation and annotation differences. The proposed comparison concerns tumor niches, immune exhaustion, treatment timing; it would test whether the specific failure scenario is detected without erasing legitimate uncertainty. An integrated account should preserve the distinction between tissue contexts, clinical endpoints, and mechanistic claims and any stronger downstream decision.

Why this distinction matters

Tumor niches and treatment timing occupy different positions in the evidence chain. Their relationship is the starting point for examining spatiotemporal sampling. Consider the possibility of a biopsy snapshot presented as a stable tumor-wide state. This is a proposed stress scenario, not an observed event reported by the cited studies. Its purpose is to expose which evidence contract would be violated even if the final output looked plausible. For tissue contexts, clinical endpoints, and mechanistic claims, the relevant claim must identify its population or artifact class, its time window, and the action it is supposed to support. A failure at any of these levels may persist even after a local technical improvement. The analysis therefore treats published methods as bounded evidence and asks how their limitations interact with this particular decision. An operational analogy is useful only if it keeps those limitations visible.

A spatial snapshot is a sample of a changing process. Apparent niches can reflect sampling location, treatment timing, or transient cell states. Joint spatial and temporal design is needed before interpreting a pattern as stable organization. In this article, the relevant contrast is tumor niches, immune exhaustion, treatment timing. The proposed counterexample of a biopsy snapshot presented as a stable tumor-wide state puts that contrast under pressure. Retain sampling coordinates and treatment chronology. Compare patterns across times and regions, with controls for preparation and annotation differences. A favorable finding would show that the intended interpretation of treatment timing remains defensible when tumor niches is perturbed under the specified conditions. An unfavorable finding should narrow the claim rather than be discarded as an inconvenient case.

The strongest comparison would isolate the disputed transition in tumor niches, immune exhaustion, treatment timing. The proposed counterexample, a biopsy snapshot presented as a stable tumor-wide state, identifies what should not be accepted as equivalent evidence. The input record should retain enough information to decide which part of that scenario violates the intended contract. That includes the original source type, measurement or artifact status, and the reason the proposed interpretation exceeds its support. The evaluator would compare the system's response with a qualified judgment made without seeing that response. Disagreement should be classified rather than erased by rewriting the prompt. Joint spatial and temporal design is needed before interpreting a pattern as stable organization. A useful negative result would show that the interface cannot preserve this distinction under realistic use; a useful positive result would

demonstrate preservation on independent cases rather than merely on the example used to develop the rule. The primary source remains relevant through its particular contribution: The set-point framing asks whether durable remission requires changing a coupled metabolic state rather than only lowering an inflammatory marker. That contribution should be retained even if the integrated workflow fails. Conversely, successful integration would not remove the source's limitations or turn a proposed extension into an already observed result. This asymmetric interpretation protects both the literature and the user from an unjustified all-or-nothing verdict.

Evidence roles and limits

(Xiao Yu et al. 2026e) examines why autoimmune activity can persist or recur when conventional inflammation suppression does not durably change underlying cellular and tissue states in *Resetting Immunometabolic Set Points in Autoimmune Disease*. Its methodological basis is a conceptual review proposing immunometabolic set points across immune and tissue compartments. The accessible evidence is specific: the review integrates cellular metabolism, nutrient/oxygen conditions, mitochondrial stress, and stromal metabolites as candidate mechanisms of persistence. The set-point framing asks whether durable remission requires changing a coupled metabolic state rather than only lowering an inflammatory marker. A unifying set point is a hypothesis, not an already validated clinical measurement or an established therapeutic protocol. The source supplies one part of the comparison involving tumor niches, immune exhaustion, treatment timing. Its evidence should retain its original unit and scope; the present synthesis proposes a connection without claiming that the paper directly tested this integrated workflow.

A useful boundary case is supplied by *Immune Niches in Brain Tumors: Glial-Immune Cell Interactions and Spatial Microdomains*, concerned with how spatial glial-immune interactions organize permissive and suppressive microdomains within brain tumors (Qian Guo et al. 2026). Its methodological basis is a review of astrocytes, microglia, oligodendrocyte-lineage cells, T cells, and myeloid cells across spatial tumor niches. The accessible evidence is specific: the abstract integrates cytokine, metabolic, hypoxic, and immunoediting mechanisms associated with immune suppression and exhaustion. It shifts attention from tumor-wide averages to interactions whose meaning depends on location and cell context. The review does not establish efficacy of a new intervention; co-location, state annotation, and functional interaction require different evidence. The contribution becomes relevant to spatiotemporal sampling when the workflow moves from tumor niches to treatment timing. That transition is the review's proposed analytical focus, not a new result attributed to the source.

A useful boundary case is supplied by *Splenic Extramedullary Granulopoiesis and Neutrophil States in Cancer Immunity: Systemic Immunosuppression and Therapeutic Resistance*, concerned with how splenic extramedullary granulopoiesis and neutrophil heterogeneity contribute to systemic cancer immunosuppression and resistance (Yidan Zhang et al. 2026). Its methodological basis is a review organized around stress-associated splenic blood-cell production and reprogrammed neutrophil states. The accessible evidence is specific: the official Cancer Bioscience recent-articles listing confirms the review's scope and emphasizes sustained stress and systemic reprogramming; no new patient cohort is inferred. The framing connects extra-tumoral immune production with local cancer resistance questions. The accessible summary cannot establish specific quantitative contributions, universal cell-state functions, or clinical effects of proposed targeting strategies. Reading the work through spatiotemporal sampling changes the evaluation question: the issue is whether tumor niches supports the intended interpretation of treatment timing, rather than whether the reported method is impressive in its original task.

(Bibo Gao et al. 2026) examines which retrospective clinical features predict extended hospital stay after endoscopic endonasal surgery for nonfunctioning pituitary adenomas in *The Prediction of Extended Hospital*

Length of Stay in Patients after Endoscopic Endonasal Transsphenoidal Surgery for the Resection of Non-Functioning Pituitary Adenomas: A Dual-Center Retrospective Analysis. Its methodological basis is a dual-center retrospective prediction analysis with a modeled extended-stay outcome and separate validation. The accessible evidence is specific: the accessible record describes a nomogram and a 100-patient validation sample; specific predictor effects and thresholds are not assumed. It treats hospitalization burden as an explicitly predicted outcome that can inform care-process evaluation. Length of stay includes organizational and discharge factors; it is not an interchangeable measure of tumor biology, surgical quality, or long-term recovery. The source supplies one part of the comparison involving tumor niches, immune exhaustion, treatment timing. Its evidence should retain its original unit and scope; the present synthesis proposes a connection without claiming that the paper directly tested this integrated workflow.

(Yinghao. “Towards Explainable RAG: Interpreting the Influence of Retrieved Passages on Generation.” *2025 4th International Conference on Robotics Sang 2025) examines how generated answers can be interpreted through the influence of their retrieved passages in Towards Explainable RAG: Interpreting the Influence of Retrieved Passages on Generation. Its methodological basis is the supplied IEEE conference title frames passage-influence analysis for explainable retrieval-augmented generation; implementation particulars were unavailable. The accessible evidence is specific: the bibliographic item is located in the official IEEE record; no unverified attribution algorithm or benchmark result is inferred. It asks for traceability between particular retrieved evidence and generated statements instead of treating a bibliography as an explanation. Passage influence, citation support, causal explanation, and factual correctness are distinct properties; the accessible record cannot establish all of them. This source enters the present review as a methodological comparison for evidence processing, not as evidence about disease biology, patient outcomes, or clinical efficacy. The transferable question is whether tumor niches remains correctly interpreted after a computational choice involving treatment timing. Its task-specific validation cannot be borrowed to validate the clinical claim.

The sources should be compared through the roles they play, not merged into a single performance claim. Resetting Immunometabolic Set Points in Autoimmune Disease addresses why autoimmune activity can persist or recur when conventional inflammation suppression does not durably change underlying cellular and tissue states (Xiao Yu et al. 2026e). By contrast, Towards Explainable RAG: Interpreting the Influence of Retrieved Passages on Generation addresses how generated answers can be interpreted through the influence of their retrieved passages (Yinghao. “Towards Explainable RAG: Interpreting the Influence of Retrieved Passages on Generation.” *2025 4th International Conference on Robotics Sang 2025). That difference matters for spatiotemporal sampling: the evidence concerning tumor niches is not automatically evidence concerning treatment timing. A successful software operation, a plausible mechanistic account, and a validated outcome are different achievements. Their integration may improve traceability, but it cannot create an unmeasured link. A coherent review should identify where the sources complement one another and where their scopes remain incompatible. Here the proposed failure scenario is a biopsy snapshot presented as a stable tumor-wide state. The analytical value of that scenario is that it can occur without requiring any cited paper to be wrong: an otherwise valid local finding can be misused at a later interface. Preventing that misuse requires evidence about the interface itself, not simply a larger collection of local results.

Testing the interpretation

Causal inference requires a clearly specified intervention, time order, and exchangeability assumptions (Hernán and Robins 2020). A prediction can be accurate without identifying what an intervention would change. A useful causal analysis begins by defining the hypothetical comparison, eligibility, and time zero. Its assumptions concern the process generating exposure and observation, not merely the flexibility of a fitted model. In this review, causal vocabulary is appropriate only when that comparison can be specified and defended. This is methodological support for examining spatiotemporal sampling, not a claim that the source studied immune-state and tumor evidence synthesis. A separate application-specific test is required at the interface described in this review.

Epidemiologic design makes confounding, selection, and measurement error part of interpretation (Rothman, Greenland, and Lash 2008). Confounding, selection, and measurement error are distinct problems and need not move an estimate in the same direction. An analysis should describe them before deciding which adjustment is adequate. A larger dataset can reduce random error while leaving systematic error almost untouched. Precision should therefore be interpreted alongside the quality of the design and the eligibility of its population. For the comparison involving tumor niches, immune exhaustion, treatment timing, the source supplies an analytical obligation. The review's proposed application remains a hypothesis, with the deliberately difficult scenario serving as a check on scope inflation.

False-discovery control is central when many molecular features are screened (Benjamini and Hochberg 1995). Multiplicity arises from the family of questions actually searched. Controlling false discoveries within a named family does not automatically cover later tissue filters, alternative endpoints, or narrative choices. The proposed audit should retain exploratory branches so the family is visible. Independent confirmation then addresses a different question from whether a discovery threshold was correctly applied during screening. Within immune-state and tumor evidence synthesis, this distinction is relevant to spatiotemporal sampling. The proposed failure scenario is a biopsy snapshot presented as a stable tumor-wide state; its evaluation must preserve the original source's methodological scope.

Medical machine learning must fit data provenance, care pathways, and prospective evaluation (Rajkomar, Dean, and Kohane 2019). The clinical relevance of a model depends on how its output changes a care process. Development data can contain information that becomes available only after the intended decision. A review must identify that availability boundary before comparing algorithms. Prospective use also changes who is measured and who is treated, so evaluation of a frozen historical dataset cannot establish the entire future workflow. The connection to glial-immune microdomains in a longitudinal review is a constraint on interpretation. It would be inappropriate to borrow the source's methodological authority as direct confirmation of the integrated domain claim.

Responsible health AI connects prediction with workflow, monitoring, and patient impact (Wiens et al. 2019). The point at which a model reaches practice is not the end of evaluation. New populations, changing measurements, and feedback can alter failure modes. A responsible deployment needs a route for recognizing deterioration and a person authorized to respond. Monitoring should be linked to feasible corrective actions; an alert that cannot change the workflow provides accountability only in a very limited sense. The connection to glial-immune microdomains in a longitudinal review is a constraint on interpretation. It would be inappropriate to borrow the source's methodological authority as direct confirmation of the integrated domain claim.

Calibration separates discrimination from confidence used in clinical escalation (Guo et al. 2017). Discrimination and calibration deserve separate evaluation. A score can order cases well yet systematically overstate its certainty. Aggregate calibration can also hide local miscalibration near the threshold where

action changes. An audit should retain predicted probabilities and observed outcomes rather than report only correct classifications, and it should evaluate calibration on data not reused to choose the correction. An implementation should expose this limitation to reviewers of tissue contexts, clinical endpoints, and mechanistic claims. The result and its evidence qualification should remain recoverable even when the interface condenses them into a short report.

BM25 remains an important lexical baseline for retrieval systems (Robertson and Zaragoza 2009). Lexical retrieval remains a useful baseline because it is explicit about matching the available terms. Semantic matching can add recall while confusing specialized senses that share broad context. A retrieval evaluation should include both exact-identifier and meaning-based cases. Comparing a new retriever with an appropriately tuned lexical baseline helps determine whether extra computation adds support rather than merely more plausible-looking passages. The connection to glial-immune microdomains in a longitudinal review is a constraint on interpretation. It would be inappropriate to borrow the source's methodological authority as direct confirmation of the integrated domain claim.

Retrieval-augmented generation makes external evidence part of the generation process (Lewis et al. 2020). Making documents available to a generator changes its information access but does not guarantee that a statement is supported. A useful analysis separates document selection from generation conditioned on those documents. Missing evidence, irrelevant context, and selective quotation create different failures. The proposed workflow should preserve retrieved passages and evaluate claim support independently of fluency or the mere presence of citations. This is methodological support for examining spatiotemporal sampling, not a claim that the source studied immune-state and tumor evidence synthesis. A separate application-specific test is required at the interface described in this review.

Dense retrieval provides semantic matching but must be compared with strong lexical baselines (Karpukhin et al. 2020). Semantic retrieval can connect questions and passages whose wording differs. That advantage should be tested against cases where wording differs because the underlying population or mechanism differs. Retrieval accuracy is not identical to answer accuracy, and neither automatically establishes evidence eligibility. A good evaluation retains the ranked candidates so later errors can be localized to retrieval rather than attributed to the generator by default. The connection to glial-immune microdomains in a longitudinal review is a constraint on interpretation. It would be inappropriate to borrow the source's methodological authority as direct confirmation of the integrated domain claim.

Replication and transparent workflows are prerequisites for reliable biomedical claims (National Academies 2019). Reproducing a calculation and independently replicating a finding are different achievements. Recoverable data and code allow one to inspect the first; independent observations and design choices inform the second. A review should avoid calling several reanalyses of one dataset separate replications. Clear reporting of analysis changes also makes a failed replication informative instead of allowing it to disappear as an implementation dispute. The connection to glial-immune microdomains in a longitudinal review is a constraint on interpretation. It would be inappropriate to borrow the source's methodological authority as direct confirmation of the integrated domain claim.

Proposed comparisons and controls

Retain sampling coordinates and treatment chronology. Compare patterns across times and regions, with controls for preparation and annotation differences. This would operationalize the concern with spatiotemporal sampling rather than assume it has been resolved by one successful output. These procedures are proposed, not completed. The case set would include the scenario of a biopsy snapshot presented as a stable tumor-wide state, with its validity judged independently of the system's response. The relevant comparison axes are tumor niches, immune exhaustion, treatment timing.

A discriminating protocol would specify a preferred interpretation and at least one credible alternative before choosing measurements. It would include negative controls and observations that could favor the alternative. Reviewers would classify whether disagreement concerns data quality, model structure, or the scientific question itself. The comparison should reduce uncertainty about that distinction rather than add more observations equally compatible with every account.

For the immune review, proposed mechanisms should be tested in their stated compartments and disease settings. Spatial association, state annotation, and intervention response would be judged separately. A useful result could expose a missing bridge in the integrated account. It would not need to confirm every part of the preferred narrative to improve scientific understanding.

Limits of the synthesis

Access limits are material for Towards Explainable RAG: Interpreting the Influence of Retrieved Passages on Generation. Where only bibliographic scope or a limited abstract is available, the discussion does not assign unverified datasets, algorithms, effect sizes, or clinical recommendations. This review is a bounded conceptual synthesis rather than a systematic review with an exhaustive search protocol. It cannot establish the effectiveness of the proposed integrated workflow. Its specific contribution is to articulate how spatiotemporal sampling changes the interpretation of tumor niches, immune exhaustion, treatment timing. Technical sources support computational comparisons only; they do not validate biological mechanisms or patient benefit. Reviews and correspondence retain their publication roles and are not counted as new experiments. Proposed stress cases are illustrative, not observed failures attributed to the papers. Future work should therefore test the weakest inference described here, preserve contrary results, and disclose the cost of checking rather than treating additional automation as intrinsically beneficial.

Conclusion

The most valuable next result would be a discriminating test of spatiotemporal sampling, including a realistic way for the preferred approach to fail. Retain sampling coordinates and treatment chronology. Compare patterns across times and regions, with controls for preparation and annotation differences. The remaining problem can be stated as a testable contrast involving tumor niches, immune exhaustion, treatment timing, with a biopsy snapshot presented as a stable tumor-wide state as a deliberately challenging alternative. A trustworthy workflow would preserve the evidence source, the assumptions of the transition, and the right to stop when support is insufficient. The proposed evaluation agenda gives those obligations observable tests while withholding claims of completed validation.

References

1. Yu, Xiao, et al. "Resetting Immunometabolic Set Points in Autoimmune Disease." *Journal of Translational Autoimmunity*, vol. 13, 2026, <https://doi.org/10.1016/j.jtauto.2026.100391>.
2. Guo, Qian, et al. "Immune Niches in Brain Tumors: Glial-Immune Cell Interactions and Spatial Microdomains." *Acta Neuropathologica Communications*, vol. 14, 2026, <https://doi.org/10.1186/s40478-026-02373-0>.
3. Zhang, Yidan, et al. "Splenic Extramedullary Granulopoiesis and Neutrophil States in Cancer Immunity: Systemic Immunosuppression and Therapeutic Resistance." *Cancer Bioscience*, 2026, <https://doi.org/10.1016/j.cbs.2026.100009>.
4. Gao, Bibo, et al. "The Prediction of Extended Hospital Length of Stay in Patients after Endoscopic Endonasal Transsphenoidal Surgery for the Resection of Non-Functioning Pituitary Adenomas: A Dual-Center Retrospective Analysis." *Cancers*, vol. 18, no. 10, 2026, <https://doi.org/10.3390/cancers18101582>.
5. Sang, Yinghao. "Towards Explainable RAG: Interpreting the Influence of Retrieved Passages on Generation." *2025 4th International Conference on Robotics, Artificial Intelligence and Intelligent Control (RAIIC)*, IEEE, 2025, pp. 397-400, <https://doi.org/10.1109/raic65850.2025.11170170>.
6. Hernán, Miguel A., and James M. Robins. *Causal Inference: What If*. Chapman & Hall/CRC, 2020.
7. Rothman, Kenneth J., Sander Greenland, and Timothy L. Lash. *Modern Epidemiology*. 3rd ed., Lippincott Williams & Wilkins, 2008.
8. Benjamini, Yoav, and Yosef Hochberg. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B*, vol. 57, no. 1, 1995, pp. 289-300.
9. Rajkomar, Alvin, Jeffrey Dean, and Isaac Kohane. "Machine Learning in Medicine." *New England Journal of Medicine*, vol. 380, 2019, pp. 1347-58.
10. Wiens, Jenna, et al. "Do No Harm: A Roadmap for Responsible Machine Learning for Health Care." *Nature Medicine*, vol. 25, 2019, pp. 1337-40.
11. Guo, Chuan, et al. "On Calibration of Modern Neural Networks." *Proceedings of the 34th International Conference on Machine Learning*, 2017, pp. 1321-30.
12. Robertson, Stephen, and Hugo Zaragoza. "The Probabilistic Relevance Framework: BM25 and Beyond." *Foundations and Trends in Information Retrieval*, vol. 3, no. 4, 2009, pp. 333-89.
13. Lewis, Patrick, et al. "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks." *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 9459-74.
14. Karpukhin, Vladimir, et al. "Dense Passage Retrieval for Open-Domain Question Answering." *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 2020, pp. 6769-81.
15. National Academies of Sciences, Engineering, and Medicine. *Reproducibility and Replicability in Science*. National Academies Press, 2019.