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.

Six checks before reading a spatial trend as biology: a preregistered audit of three composition paradigms on one liver cohort

SOON WOO NAM

drswnam@catholic.ac.kr

The Catholic University of Korea Incheon St. Mary's Hospital

Research Article

Keywords: spatial transcriptomics, cell type deconvolution, tissue composition, reporting checklist, liver zonation, preregistration, negative controls

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11849736/v1

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

Additional Declarations: The authors declare no competing interests.

Title: Six checks before reading a spatial trend as biology: a preregistered audit of three composition paradigms on one liver cohort

Running title: Six checks before reading a spatial trend

Author: Soon Woo Nam, MD, PhD

Affiliation: Hepatobiliary Unit, Division of Gastroenterology, Department of Internal Medicine, Incheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

ORCID: 0000-0001-5562-4775

Corresponding author: Soon Woo Nam, drswnam@catholic.ac.kr, 56 Dongsu-ro, Bupyeong-gu, Incheon 21431, Republic of Korea. Telephone +82-32-280-7493, fax +82-32-280-6010.

Author biography: Soon Woo Nam is Professor of Internal Medicine in the Division of Gastroenterology at Incheon St. Mary's Hospital, a hepatologist whose research applies preregistered reanalysis to public liver genomics data.

Abstract

A Visium spot aggregates transcripts from many cells, so a gene varying along a spatial axis may be changing activity or tracking abundance. Three routes separate them: deconvolution into per-spot fractions, a reference-derived proxy for where transcripts sit, and adjusting a spatial model for those. Benchmarks rank deconvolution against synthetic

mixtures; whether they agree on real tissue is rarely tested. Table 6 gives six checks to run before reading a spatial trend as biology; they name no tissue, and the axis and gates behind them run unchanged on a non-liver section. They are needed because the three routes disagree on one human liver cohort, under rules fixed in advance. Of nine estimator-by-reference combinations four clear the reference premise and one is refused by its estimator; on one amended arm three return 0.2357, 0.5871 and 0.7526. The lobular axis, the designed comparator, is the one the fractions track, at a median correlation of 0.7775 against 0.0925, so residualizing there is not well posed. The checks bite on this paper's own work, refuting a proxy I proposed and withdrawing a magnitude claim that component sampling cannot pin down. Every route, withdrawn ones included, is preregistered and deposited.

Keywords: spatial transcriptomics; cell type deconvolution; tissue composition; reporting checklist; liver zonation; preregistration; negative controls

Introduction

A Visium spot is not a cell. A spot measurement sums whatever lay under it [1], and in liver each 55-micrometer spot spans several cells of more than one type. A gene higher in one region is therefore ambiguous: the cells expressing it are doing more of it there, or there are more of those cells.

Three routes are in common use. The first is deconvolution: a single-cell reference supplies a profile per cell type, the mixture at each spot is estimated against those profiles, and the fractions become covariates [2–5]. The second, when deconvolution is not trusted, is a

reference-derived proxy for where a gene's transcripts sit. The third is to estimate fractions anyway and adjust the spatial model for them, as published tools do [6]. The three are usually treated as a hierarchy, fractions best and a proxy a fallback, and are rarely run side by side, because the data that would allow it are those on which deconvolution is in doubt. Benchmarks rank estimators against mixtures whose composition is known by construction [3]. The question here is one a benchmark cannot pose: on one real cohort, with no composition known, do the three routes agree, and does the reference meet the premises its use assumes.

The kind of spatial coordinate now in common use sharpens the question. A lesion-anchored coordinate, such as distance from a fibrotic septum, is correlated with composition by construction, because a lesion displaces parenchyma and recruits immune cells [7]. Composition is not an orthogonal nuisance on such an axis; it is part of what the axis is. The usual safeguard, an empirical null of genes matched on abundance, controls how much a gene is expressed and says nothing about which cells express it, so combining the two can return compositional structure as a spatial finding.

This paper runs all three routes on one human liver Visium cohort spanning the MASLD spectrum [8], against one reference [9], on two coordinates reconstructed from the same sections: the lobular axis, a metabolic pericentral to periportal coordinate running largely within the hepatocyte compartment [10, 11], and the distance axis from fibrotic tissue, a lesion coordinate crossing compartments. If the three routes agree, the lesion axis should look far more compositional than the lobular one under all of them. The comparison is the contribution.

The second route needs a name for its proxy. I call it carriage: the share of a gene's transcripts that sits in a given annotated class within a single-cell reference, pooled across the cells of that class. I proposed it as a measurement and it does not survive its own controls, so it appears here as the worked example of a proxy rather than as a method offered for use. A per-gene share across cell types is not new. The specificity metric of expression-weighted cell-type enrichment [12] normalizes each type's mean first, removing class abundance, while carriage pools raw counts, so it is a mass share of the compilation.

Every endpoint was registered before it was computed [13], each analysis is gated by premise checks that failed real data, and each pipeline carries a self-test that must fail where failure is correct. Claims made in earlier versions are withdrawn here with the deposited numbers that refute them.

Table 6 gives six checks to run before reading a spatial trend as biology, and names no tissue. The Results are the evidence, taking the three routes in order, one disagreement each: whether this reference supports a mixture fit, how compositional the two axes are, and which axis the fractions track. Two further displays matter: Table 2, the estimator-by-reference grid, and Figure 4, the collinearity that makes the third route ill posed here.

Materials and methods

Supplementary Methods gives every step below in full, with its gate values and the deposited file each number comes from.

Preregistration. Every endpoint was registered before the corresponding statistic was computed. Seven base registrations and thirty dated amendments and addenda govern the work and are deposited; each states in its own first section what had already been seen when it was written, and leaves the registered analysis in place as the primary result. Amended and exploratory quantities are labeled as such throughout.

Cohorts and units. Cohort D3 is a 10x Visium deposit of human liver covering the MASLD spectrum [8]. Because several needle biopsies sit on one capture area, the analysis unit is the tissue component, a connected component of the Visium lattice. Table S3 gives the per-area inventory; the deposit carries no donor identifier and no fibrosis stage, so the analyzed set cannot be stratified. Cohort D1 is a Visium deposit of five liver sections [9] used only for the registered replication.

Two spatial axes. The lobular axis reconstructs the pericentral to periportal coordinate, anchored on CYP2E1 and cut into ten equal-count bins, with orientation fixed per component by a landmark panel before any target gene is examined. The distance axis is the lattice-step distance from fibrotic tissue, where scar is the top quintile of a fibrosis score from a fixed 22-gene panel, so the masked fraction is 0.20 in every component whatever its collagen burden. Both are gated before use, and the components failing each gate are reported.

Trend statistic. Monotone trend is measured by the Jonckheere-Terpstra statistic [14, 15] across the bins, per component, converted to a standard normal deviate. Because at this depth most of the transcriptome gives a nominally significant trend, each target is standardized against an empirical null of control genes matched on abundance and non-

zero rate, under two matching rules that are both reported. Every null result carries its detection limit, the smallest per-unit effect the design could have resolved. Each limit stated here is the 50 percent power value; the 80 percent values recomputed under C-D10 are 1.29 to 1.56 times larger, so every null verdict understates by about that much.

The three routes. Deconvolution was run as per-spot non-negative least squares over a fixed 127-marker panel, as CARD 1.1 [4] and, under C-D14, as RCTD 2.2.1 [2], each from identical exports against the dissociated and single-nucleus arms of one liver atlas [9]. The premise check fixed before any run is stated with Table 2, and a combination that fails it is reported as not analyzed. Carriage is the share of a gene's transcripts sitting in one annotated class of the reference. For a gene g, a class c, counts $X_{g,i}$ in reference cell i, and the annotated classes $\mathcal{C}$,

$$C_{g,c} = \frac{\sum_{i \in c} X_{g,i}}{\sum_{k \in \mathcal{C}} \sum_{i \in k} X_{g,i}},$$

a mass share across classes rather than a concentration within one, so it is not invariant to class size. It enters a regression on the rank-normalized trend deviate beyond a mean-expression covariate, and the reported quantity is the increment in explained variance, $\Delta R^2 = R^2_{\text{full}} - R^2_{\text{base}}$. The composition adjustment carries the per-spot fractions into the spatial model as covariates and reports the share of gene-level trend variance removed, beside the same adjustment on shuffled fractions.

Verification. Each analysis carries a self-test that plants known structure and must fail where failure is correct. Before submission the manuscript was checked by a deposited script that binds every numeric token to a file built from the result tables, recomputes

fourteen named counts rather than all, and binds each detection limit to its endpoint. It carries 46 injected-defect controls to prove it can fail, and all 46 fired. Supplementary Methods gives what it does not cover.

Results

Two spatial axes on the same cohort, and they are not the same axis

I reconstructed two coordinates on each tissue component of cohort D3, a human liver Visium deposit spanning the MASLD spectrum [8]: a lobular axis, pericentral to periportal, and lattice-step distance from fibrosis. Both are built from data rather than annotation, both gated, and do not run on the same components: 22 pass the lobular gates, 10 the distance gates (Figure 1C, Data S4).

The two axes are weakly correlated, and the number that matters is the one from the analysis set. Across the 22 gated components the median absolute Spearman correlation is 0.227, with 2 of 22 above 0.5; across the 10 on the distance axis it is 0.245, and both components above 0.5 are inside that set, so the exposure is 2 of 10 rather than 2 of 22 (Figure 1C).

Three strongly zonated genes show the coordinates seeing different biology. SCD1, GPAM and TBX3 are inconclusive or flat in distance from scar, at standardized -1.75, -0.76 and +0.52, with detection limits of 0.714, 0.594 and 0.687 (Figure 1D).

The first disagreement: whether this reference supports a mixture fit depends on the method

CARD 1.1 [4] was the decisive test of the first route, because the original investigators ran it against this same reference. I repeated their analysis under their settings on all 22 gated components and 9,714 spots, with one departure stated in the methods, a 2,000-cell per class cap. Table 2 gives every estimator-by-reference cell against the registered premise check, which requires a median hepatocyte fraction of 0.40 or more with hepatocytes most abundant.

I first read CARD's failure as a property of the reference. It is not: least squares clears the check on the same matrix, on every component and normalization variant, with the pooled median between 0.4669 and 0.6130 (Table 2).

The obvious remedy fails under both estimators that fit it. The single-nucleus arm of the same atlas holds 0.89053 of its transcript mass in hepatocytes against 0.17016 in the compilation, yet both fail with resident NK most abundant. They agree on the reason: albumin reads 11,813.3 counts per million in resident NK nuclei against 5,381.3 in hepatocyte nuclei, so immune nuclei look parenchymal. That profile rests on almost nothing: five of this subset's ten classes hold eleven nuclei or fewer, resident NK among them (Data S5), which is why deconvolution tools impose a minimum cells per class [2]. Dropping the five under C-D4 raises the pooled median but leaves fibroblasts most abundant, so under least squares that arm still fails and the quality defect is not what makes it fail [16]. That variant passes under RCTD at 0.7526 and CARD at 0.5871, while RCTD declines the unmodified arm outright (C-D14, C-D16, Table 2). On one reference and

one set of spots the three return 0.2357, 0.5871 and 0.7526, and least squares alone fails under every variant. The defect is what one estimator cannot see past, what another passes once the arm is repaired, and what the third refuses to fit.

Estimator, reference and quality are not separable here: the reference a reader would call better fails both estimators that fit it and is refused by the third, while the one called unusable supports two. Among the four primary combinations the threshold does little work on the three that fail: across 0.25 to 0.55 none changes verdict, because each fails on which class is most abundant, which no threshold on parenchyma repairs, and two fail on parenchyma too, at 0.0950 and 0.0000 (Data S5). It does work on one added later: RCTD on the dissociated compilation sits at 0.4445 and fails at 0.45 (Table 2).

What the first route can report about the reference is the compilation's composition, the background every carriage value is read against: hepatocytes hold 0.17016 of its transcript mass against 0.33139 for T cells, a property of the atlas rather than of liver (Supplementary Results, Table 1).

The second disagreement: the proxy and the fractions agree on the direction and not on the size

The second route replaces fractions with carriage. Per axis I regressed the rank-normalized trend deviate of every eligible gene on its carriage beyond a mean-expression covariate, taking the increment in explained variance on the 11,123 shared genes. The contrast is conditional on a gate made on a quantity the regression then predicts, and under C-D13 that selection is not what produces it (Table 3).

Carriage accounts for 0.15449 of that variance on the distance axis, 0.00934 on the lobular (Figure 3A). The difference is 0.14516, with a 2,000-draw bootstrap interval of 0.13212 to 0.15813 that excludes zero (Figure 3B), confirming the registered endpoint. The ratio between axes is 16.5, and under C-D11 it is not stable: over 500 random ten-component subsets of the lobular axis it runs from 5.2 to 109.1 for 95 percent of draws, with the ordering never reversed (Table 3).

The third route can now run, the first having cleared the premise check. That licenses the hepatocyte column alone: the same fit puts cholangiocytes at 0.1906 and every lymphocyte class at zero, which no liver allows, and the check inspects neither (C-D18, Table 2). The registered composition adjustment, not analyzed since the original registration, carries per-spot fractions into the model as covariates: per gene and per component the trend is the t statistic of the axis coefficient with and without the fraction columns, combined over components as before.

Composition removes 0.7414 of the gene-level trend variance on the distance axis and 0.6065 on the lobular axis (Figure 3C), a ratio of 1.22. The axes do not run on the same components, 10 against 22. Matching them changes both routes: the lobular carriage analysis rerun on those 10 gives 16.1 under C-D8, and the lobular fractions recombined over the same 10 give 0.4945 against 0.7410, a ratio of 1.50, under amendment C-D6. The matched comparison is 16.1 against 1.50, and neither figure is separable from which ten components it is measured on.

That comparison is weaker than it looks. Resampling the components under C-D9 puts the implied ratio between 1.06 and 1.65 for 95% of draws, with 2 of 500 falling to 0.997, and

the observed 1.50 at their 90th percentile (Table 3). On the difference scale they disagree too: matched, the fractions remove 0.2465 against the proxy's 0.14241. Neither axis is the clean comparator the design assumed, and with both ratios dependent on the component set the comparison supports a disagreement, not a number for how much. The registered controls on the adjustment hold (Supplementary Results).

The third disagreement: the axis the fractions track is the lobular one

The design assumed the lesion axis compositional and the lobular axis the within-compartment comparator. Against the estimated fractions it is the reverse. The hepatocyte fraction tracks the lobular axis at a median Spearman correlation of 0.7775 across the 22 components (Figure 4A), while on the distance axis that correlation has a median of 0.0925 (Figure 4A).

On the lobular axis the correlation is positive in all 22 components, so the estimated hepatocyte share rises toward the periportal pole, which anatomy predicts the opposite of, and the reference cannot distinguish the poles, carrying CYP2E1 at 0.9673 and GLUL at 0.1609 though both are pericentral and a human atlas resolves them [17]. Supplementary Results gives three accounts these data cannot separate, including zone 3 steatosis lowering hepatocyte transcript density per spot; the consequence below is the same under each.

Adjusting a coordinate for a covariate that nearly is that coordinate does not adjust the trend, it removes it and leaves a reversed remainder. My first implementation residualized instead of adjusting, and inverted (Figure 4B). On an axis this collinear with composition, adjusting for composition is not a well-posed operation; the collinearity itself is.

The carriage association survives the nuisance-model attacks I could mount; the landmark control, reported as absent in an earlier version, meets its registered expectation in the amended model (Figure S1B, Supplementary Results). Six limits bound what the association shows, none removable with the data in hand, and Table 3 gives each with its measuring quantity. None of this supports reading the contrast as a measurement of how much more compositional the lesion axis is.

A worked pair, and what two restrictions do to it

A route is worth having if it separates a spatial result about composition from one that is not, and this dataset supplies both cases.

MTARC1 is the first. The redox panel is MTARC1, MTARC2, CYB5R3, CYB5A and POR, registered as distance-axis endpoints expected to be flat because MTARC1 p.A165T protects against steatotic liver disease and cirrhosis without a known spatial correlate [18], and read conventionally MTARC1 is a finding (Table 4). Neither of two checks supports it: none of the five redox genes reaches a residual shortlist (Figure 5B), and capture-area clustering, the ten components coming from six areas and three from one, takes it to inconclusive (Supplementary Results).

The two gates testing whether a top-quintile mask is scar were written for the replication cohort and never run here (Table 1), so the mask rests on the fibrosis score alone.

CDKN1A is the contrast that makes the caution usable. It moves away from scar while its immune carriage of 0.9367 is only 1.17 times an immune background of 0.80053, against a

ceiling of 1.25, so composition predicts the opposite direction, and it survives both restrictions (Table 4).

The procedure run on a tissue it was not written for

Under transfer registration T1 the distance-axis procedure was run unchanged on one public human prefrontal cortex Visium section [19], with the four tissue slots of Table 5 refilled and no threshold or rule altered. Every gate passed, and 3 of 8 analyzable layer markers cleared their null (Table 5). Three of the five that did not are the point: RASGRF2 at +8.5491, CARTPT at +9.1276 and TLE4 at -5.4393 run in the registered direction yet do not clear, their matched peers reaching as far.

What the gates refused, and what registered predictions got wrong

Nine outcomes here are negative, self-correcting, or contrary to what I registered. Table 1 gives each with the numbers that settle it, and Data S3 every endpoint with its verdict. One deserves more than the table.

The distance axis was not analyzable as registered, for a geometric reason: MASLD fibrosis is distributed rather than consolidated [20], so almost no parenchyma in a needle biopsy lies over three lattice steps from scar. Amendment A-F1, written before any contrast statistic, lowered the size requirement and rebinned the axis without relaxing F2c. The registered replication stopped at its premise check, so I claim nothing about replication; its two passing sections ran anyway and MTARC1 is flat there, on two units, detection limit 1.585.

The post hoc Z3 explanation is neither supported nor refuted, and the registered model omits an endothelial term on both axes (Table 1).

Discussion

Three routes to one question, on one cohort, disagree. CARD says this reference cannot be used; least squares and RCTD on the identical matrix say it can; on the amended arm the three spread from 0.2357 to 0.7526 and least squares alone fails. The proxy says a lesion-anchored axis is far more compositional than a lobular one, at 16.1 on matched components; the fractions say 1.50. Neither magnitude survives component resampling, though the proxy's ordering holds in every draw and the fractions' falls to 0.997 in 2 of 500. And the axis the design treated as the comparator is the one the fractions track, at 0.7775 against 0.0925. A paper running one route would have reported each as a finding.

The disagreements differ in kind. The first is about estimator and reference together. The second is about what is measured: carriage is a property of a compilation and estimated fractions of the tissue, and a mass share inherits the class proportions of whatever was sequenced. The third is about the coordinates, and it is the one I did not anticipate.

No route is wrong in general, but where composition and the coordinate are entangled the method determines the answer, so reporting one reports a choice as a result. A paper that estimates fractions and adjusts owes the premise check that would have failed; one using a proxy owes the comparison against fractions where those can be estimated; one on any axis owes the coordinate-covariate collinearity, because which axis carries it is not predictable, and my own first attempt assumed otherwise.

Two further habits follow. Where a spatial axis is correlated with composition, a result standardized only against an abundance-matched null is not yet one: that null controls how much a gene is expressed, not which cells express it. MTARC1 is the worked case: it cleared its null and would have been reported as a spatial finding, and two further checks fail to support it. CDKN1A reaches the shortlist and survives both restrictions. And components from one capture area are not independent: ten from six areas, analyzed as ten, turn MTARC1 and MTARC2 positive and as six leave them inconclusive. No donor identifiers are deposited, so the true number of units may be below six. Table 6 collects these checks.

On carriage the honest summary is narrow. A single threshold reproduces 0.885 of the registered distance increment, so its continuous structure is not what the contrast rests on, and under C-D5 a curated list beats it at equal flag count. Cell-type specificity, which carriage reweights by class transcript totals, predicts both axes better and halves the axis ratio, so the reweighting inflates the contrast rather than revealing it. The contrast rests on a wide hepatocyte flag, which carriage is one way to draw. The fraction-adjusted analysis is the better posed answer to the question I proposed carriage for, though on these axes it gives no identified magnitude either.

The approach needs an annotated reference, a coordinate and a trend statistic. The schema is general; the implementation names liver at six points (Table 3); Table 5 runs the axis and the gates on cortex with the four axis and gate slots refilled. Higher-resolution platforms narrow the mixture without removing the question, since a reference still says which cells a transcript belongs to [21, 22].

Limitations

The three routes are compared on one cohort and two coordinates, deconvolution against three reference variants and the others against one; the comparison is the result, not a benchmark. This paper does not adjudicate which estimator is nearer the truth. CARD is here because the depositors used it on these sections, so what fails the check is a published pipeline. RCTD [2], added under C-D14, changed which combinations clear the check: a platform-effect model recovers parenchyma from the amended single-nucleus arm, and under C-D16 CARD does too. cell2location [5] was not run. C-SIDE [6], the third route's implementation, declined all ten distance-axis components under C-D15 at its own default thresholds. No threshold was lowered, its self-test did not pass, and the registration halts a run whose self-test does not, so nothing from it is reported (Supplementary Results). What the comparison compares is an estimator together with the gene set it reads. On identical spots and the same reference export, three estimators return 0.2357, 0.5871 and 0.7526, but least squares is restricted to the 123 present genes of the registered 127-marker panel while the other two read the whole shared transcriptome. The estimator that fails the amended arm is also the one reading the smallest gene set, and this design cannot separate those. Nor do the three report the same quantity: least squares returns transcript-mass shares, CARD cell-type proportions and RCTD transcript-attributable weights, and in liver, where hepatocyte transcript content exceeds a lymphocyte's by an order of magnitude, one threshold of 0.40 does not mean the same thing across them. A check gates one estimator-by-reference pair, so clearing it says nothing about the pairs that fail. The unmodified arm fails under both estimators that fit it and is refused by the third; dropping the five classes

built on eleven nuclei or fewer takes CARD from 0.0950 to 0.5871 and makes the arm fittable for RCTD, which returned no number on it before and so cannot be said to have recovered from anything. Least squares still fails under every variant. The references also differ in size, 11,082 cells against 156,516 before the cap, confounding that contrast with depth.

Carriage is a property of the reference, not of tissue, and does not say a cell type expresses a gene in a spot. As a mass share it inherits the compilation's class proportions, so its values describe this reference rather than liver, a residual is a shortlist only, and the contrast is an association rather than an identified effect (Table 3).

The fractions carry error the adjusted trends treat as zero. The hepatocyte class under-recovers, as it does in dissociated human liver generally [23], so every hepatocyte fraction reads better as a floor than an estimate, and the adjustment requires every class in the model to be in the reference.

The two axes do not run on the same components, 22 lobular and 10 distance from six capture areas. The registered replication stopped at its premise check, so the distance-axis findings are labeled "not tested independently". Every distance-axis number comes from an amendment, labeled as such throughout. Both matching rules are reported and disagree for SCD1, where the governing rule rests on two components.

The detection limits are optimistic. The transfer in Table 5 covers the axis and the gates, not the deconvolution steps, and one section is one unit.

Conclusion

Table 6 states six things I would want to see in any paper reading a spatial trend as regulation: the reference premise tested on the dominant class and on the non-dominant classes it never inspects, the coordinate-composition collinearity, a reference-specificity residual, a null clustered at the sampling level, and a magnitude reported as an interval or not at all. None needs new data. They matter because on one cohort the three routes disagree, and which is run decides what is reported as biology. They are not advice I escaped: they refuted a proxy I proposed and withdrew a magnitude I had claimed. Strong zonation and albumin landing in the wrong class made these failures legible; a tissue with neither would fail the same checks more quietly.

Key Points

- Table 6 gives six checks to run before reading a spatial trend as biology; they name no tissue and need no new data.
- They are needed because three routes for handling tissue composition disagree on one human liver Visium cohort, so the route chosen decides the result.
- Of nine estimator-by-reference combinations four clear a premise check fixed in advance, four fail and one is refused outright by its estimator; on one amended arm three estimators return hepatocyte medians of 0.2357, 0.5871 and 0.7526.
- The axis assumed to be the within-compartment comparator is the one the fractions track, at a median Spearman correlation of 0.7775 against 0.0925.

- The checks bite here: a proxy I proposed fails its own controls, and clearing a premise check licenses the dominant class alone, since the four that clear it lead with four different non-hepatocyte classes.
- Every endpoint was registered before the corresponding statistic was computed, and the deposited verifier carries 46 defect controls.

Acknowledgments

This work used public data deposited by other investigators, and I thank them for making it available.

Funding

No external funding supported this study.

Author contributions

S.W.N. is the sole author and is responsible for the conception, the preregistrations, the analysis, the figures, and the writing.

Declaration of interests

The author declares no competing interests.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the author used Claude (Anthropic) to carry out literature searches, to prepare a first draft of the text, and to implement the statistical analyses and figures specified by the author. The concept, the hypotheses, the study design and all preregistered decision rules are the author's own. The author reviewed and edited the content and takes full responsibility for it.

Availability of data and materials

This study analyzes existing public data and generated no new materials. The primary datasets, with their accessions and DOIs, are in Supplementary Table S1. All original code, the seven preregistrations and the thirty dated amendments and addenda, and every result table from which a number in this paper is taken, are deposited at Zenodo, https://doi.org/10.5281/zenodo.24096240, and are public without registration.

References

References 1 to 23 are cited in the main text in order of first citation. References 24 to 27 are cited only in the Supplementary material and are listed here so that one numbering serves both files.

1. Ståhl PL, Salmén F, Vickovic S et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science 2016;353:78–82. doi: 10.1126/science.aaf2403.
2. Cable DM, Murray E, Zou LS et al. Robust decomposition of cell type mixtures in spatial transcriptomics. Nat Biotechnol 2022;40:517–26. doi: 10.1038/s41587-021-00830-w.
3. Li H, Zhou J, Li Z et al. A comprehensive benchmarking with practical guidelines for cellular deconvolution of spatial transcriptomics. Nat Commun 2023;14:1548. doi: 10.1038/s41467-023-37168-7.
4. Ma Y, Zhou X. Spatially informed cell-type deconvolution for spatial transcriptomics. Nat Biotechnol 2022;40:1349–59. doi: 10.1038/s41587-022-01273-7.
5. Kleshchevnikov V, Shmatko A, Dann E et al. Cell2location maps fine-grained cell types in spatial transcriptomics. Nat Biotechnol 2022;40:661–71. doi: 10.1038/s41587-021-01139-4.
6. Cable DM, Murray E, Shanmugam V et al. Cell type-specific inference of differential expression in spatial transcriptomics. Nat Methods 2022;19:1076–87. doi: 10.1038/s41592-022-01575-3.
7. Ramachandran P, Dobie R, Wilson-Kanamori JR et al. Resolving the fibrotic niche of human liver cirrhosis at single-cell level. Nature 2019;575:512–18. doi: 10.1038/s41586-019-1631-3.
8. Vu H, Sun Y, Xiong Z et al. Progressive fibrosis in human MASLD is associated with spatially linked transcriptomic signatures of metabolic reprogramming and senescence. JHEP Rep 2026;8:101657. doi: 10.1016/j.jhepr.2025.101657.

9. Guilliams M, Bonnardel J, Haest B et al. Spatial proteogenomics reveals distinct and evolutionarily conserved hepatic macrophage niches. Cell 2022;185:379–396.e38. doi: 10.1016/j.cell.2021.12.018.
10. Halpern KB, Shenhav R, Matcovitch-Natan O et al. Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. Nature 2017;542:352–6. doi: 10.1038/nature21065.
11. Ben-Moshe S, Itzkovitz S. Spatial heterogeneity in the mammalian liver. Nat Rev Gastroenterol Hepatol 2019;16:395–410. doi: 10.1038/s41575-019-0134-x.
12. Skene NG, Grant SGN. Identification of vulnerable cell types in major brain disorders using single cell transcriptomes and expression weighted cell type enrichment. Front Neurosci 2016;10:16. doi: 10.3389/fnins.2016.00016.
13. Nosek BA, Ebersole CR, DeHaven AC et al. The preregistration revolution. Proc Natl Acad Sci USA 2018;115:2600–6. doi: 10.1073/pnas.1708274114.
14. Jonckheere AR. A distribution-free k-sample test against ordered alternatives. Biometrika 1954;41:133–45. doi: 10.1093/biomet/41.1-2.133.
15. Terpstra TJ. The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. Indag Math 1952;55:327–33. doi: 10.1016/S1385-7258(52)50043-X.
16. Andrews TS, Atif J, Liu JC et al. Single-cell, single-nucleus, and spatial RNA sequencing of the human liver identifies cholangiocyte and mesenchymal heterogeneity. Hepatol Commun 2022;6:821–40. doi: 10.1002/hep4.1854.
17. Aizarani N, Saviano A, Sagar et al. A human liver cell atlas reveals heterogeneity and epithelial progenitors. Nature 2019;572:199–204. doi: 10.1038/s41586-019-1373-2.

18. Emdin CA, Haas ME, Khera AV et al. A missense variant in Mitochondrial Amidoxime Reducing Component 1 gene and protection against liver disease. PLoS Genet 2020;16:e1008629. doi: 10.1371/journal.pgen.1008629.
19. Maynard KR, Collado-Torres L, Weber LM et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. Nat Neurosci 2021;24:425–36. doi: 10.1038/s41593-020-00787-0.
20. Kleiner DE, Brunt EM, Van Natta M et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313–21. doi: 10.1002/hep.20701.
21. Oliveira MF, Romero JP, Chung M et al. High-definition spatial transcriptomic profiling of immune cell populations in colorectal cancer. Nat Genet 2025;57:1512–23. doi: 10.1038/s41588-025-02193-3.
22. Janesick A, Shelansky R, Gottscho AD et al. High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. Nat Commun 2023;14:8353. doi: 10.1038/s41467-023-43458-x.
23. MacParland SA, Liu JC, Ma X-Z et al. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. Nat Commun 2018;9:4383. doi: 10.1038/s41467-018-06318-7.
24. Fujiwara N, Kubota N, Crouchet E et al. Molecular signatures of long-term hepatocellular carcinoma risk in nonalcoholic fatty liver disease. Sci Transl Med 2022;14:eabo4474. doi: 10.1126/scitranslmed.abo4474.
25. Moran PAP. Notes on continuous stochastic phenomena. Biometrika 1950;37:17–23. doi: 10.1093/biomet/37.1-2.17.

26. Svensson V. Droplet scRNA-seq is not zero-inflated. Nat Biotechnol 2020;38:147–50. doi: 10.1038/s41587-019-0379-5.

27. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 1995;57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x.

Figure titles and legends

Figure 1. Two spatial axes reconstructed on the same tissue, and the discriminant illustration that separates them. (A) One tissue component (VLP115_A_c2, 423 spots), colored by position on the lobular axis. (B) The same component colored by lattice-step distance from fibrotic tissue, scar spots ringed as the top quintile of the fibrosis score. The amended axis bins distances as 0, 1, 2, 3 and 4 or more. (C) Absolute Spearman correlation between the two axes across the 22 gated components. Solid line, the median of 0.227; dashed line, the registered threshold of 0.5. Across the 10 distance-axis components the median is 0.245. (D) Three strongly zonated genes on both axes, as combined standardized statistics under the quantile rule, with rules at the 97.5th percentile of each gene's abundance-matched null. SCD1 is the lobular positive control, here on 22 components; its governing fixed-rule lobular value is +4.7579 on 6. GPAM and TBX3 were designated exploratory. All three are strongly pericentral; on the distance axis GPAM and TBX3 are flat and SCD1 inconclusive. They rank 10th, 35th and 90th on the deposited lobular ranking.

Alt text: Four panels. Panels A and B show the same tissue component as a honeycomb of Visium spots, colored first by position on the lobular axis from pericentral to periportal and

then by lattice-step distance from scar, with scar spots ringed. Panel C is a histogram of the absolute Spearman correlation between the two axes across the 22 gated components, with a line at the median of 0.227 and a dashed line at the separability threshold of 0.5. Panel D is a grouped bar chart of three strongly zonated genes, each with a tall bar on the lobular axis and a short bar near zero on the distance axis.

Figure 2. Whether a reference supports a mixture fit depends on the estimator as well as the reference. (A) Median CARD-estimated proportion across 9,714 spots in 22 components, for classes reaching 0.005. The registered premise required hepatocytes to reach 0.40 and be most abundant; they reach 0.3137 while fibroblasts lead at 0.3937. (B) Pooled median hepatocyte fraction for all nine estimator-by-reference combinations under the same registered premise check: per-spot non-negative least squares, CARD 1.1 and RCTD 2.2.1, each against the dissociated 17-class reference, the single-nucleus 10-class subset of the same atlas and the amended variant of that subset, on the same 9,714 spots and 22 components. Solid outline marks a combination that passes; each bar's label gives the verdict and the components passing. The dashed outline is the one cell no estimator fitted: RCTD refuses the unmodified single-nucleus subset, so it returns no fraction to check. The panel shows only the first limb of the check; Table 2 adds each cell's leading non-hepatocyte class and immune total, which the check does not inspect.

Alt text: Two panels. Panel A is a horizontal bar chart of median CARD proportions by cell class across 9,714 spots, with fibroblasts the longest bar at 0.3937, hepatocytes next at 0.3137, and a dashed vertical line marking the registered premise of 0.40. Panel B is a grouped bar chart of the pooled median hepatocyte fraction for all nine estimator-by-

reference combinations, three estimators against three reference variants, with a dashed horizontal line at the registered premise of 0.40. Four bars rise above it and are labeled as passing, four fall below it and are labeled as failing, and one position carries a dashed empty outline labeled refused.

Figure 3. The reference-derived proxy and the estimated fractions make the lesion axis the more compositional one and disagree on by how much. (A) The proxy. Increment in explained variance from the carriage terms beyond a mean-expression covariate, per axis, under the registered two-term model and under amendment A-C1 with five terms, which is secondary throughout. Each bar carries its own permutation null at the 97.5th percentile, 0.00062 to 0.00122. The ratio between axes is 16.5. Every distance-axis quantity here is an amendment A-F1 result; the bar labels distinguish the A-C1 carriage models, not the axis definition. (B) The registered primary endpoint, the difference between axes, on 11,123 shared genes, with 2,000-draw bootstrap intervals over genes only. (C) The fractions. Share of gene-level trend variance removed by carrying the per-spot fractions into the spatial model as covariates, per axis, beside the same adjustment on fractions shuffled within component. The ratio between axes is 1.22 as the axes are run and 1.50 on matched components under amendment C-D6, where random ten-component subsets of the lobular axis give 1.06 to 1.65 in 95% of draws (amendment C-D9). Each share is bounded by one, so the numerator of this ratio is bounded where the proxy's is not.

Alt text: Three panels. Panel A is a bar chart of the carriage increment in explained variance on each axis under the registered two-term model and the amended five-term model, with the distance-axis bars far taller than the lobular ones and a dashed line marking the

permutation null. Panel B is a point-and-interval plot of the primary endpoint, the difference between axes, under the registered and amended models, with bootstrap intervals that exclude zero. Panel C is a bar chart of the share of gene-level trend variance removed by the estimated fractions on each axis, each beside a much shorter bar for fractions shuffled within component.

Figure 4. The axis the fractions track is the lobular one, and a single threshold on the proxy reproduces the distance-axis contrast. (A) Spearman correlation between the estimated hepatocyte fraction and the axis, one point per component, for each axis. Rules are medians. The fractions track the lobular axis at a median of 0.778 and the distance axis at 0.092, the reverse of what the design assumed. On the lobular axis the correlation is positive in all 22 components and never below 0.5182, and the fraction matrix explains a median 0.6154 of the zone index; on the distance axis the correlation runs from -0.2944 to 0.5794 and the matrix explains a median 0.3386. (B) Distance-axis increment in explained variance when the two registered carriage terms are replaced by coarser substitutes, on exactly the registered gene set with the registered response and covariate. A binary indicator at a hepatocyte carriage of 0.80 keeps 0.885 of the registered increment, while membership of a fixed 16-gene hepatocyte marker panel, which reads no carriage value, keeps two percent. The first implementation, which residualized expression on the fractions and tested the residuals for trend, returned adjusted trends 3.9 times as variable as unadjusted ones on the lobular axis, anti-correlated with them at Spearman -0.562, while its permuted control removed 0.003.

Alt text: Two panels. Panel A is a strip plot of the Spearman correlation between the estimated hepatocyte fraction and the axis, one point per component, for each of the two axes, with the lobular points clustered high around a median of 0.778 and the distance points scattered around 0.092. Panel B is a horizontal bar chart of the distance-axis increment in explained variance when the two carriage terms are replaced by coarser substitutes, with the registered model longest, the binary and tertile substitutes close behind, and the 16-gene marker panel a very short bar.

Figure 5. A worked pair: a finding that dissolves and a finding that holds. (A) Distance-axis combined standardized statistics for the five redox-node genes and CDKN1A against their abundance-matched nulls under the quantile rule, at the 97.5th percentile. Open markers give the capture-area clustered value with its own clustered null: MTARC1 becomes inconclusive at +1.7816 against 1.9181 while CDKN1A holds at +2.7627 against 1.8336. (B) Standardized carriage residuals for the same genes under both models. Dashed lines mark 1.8233, the smallest absolute residual reaching either distance shortlist. None of the five redox genes approaches it; CDKN1A reaches it in both.

Alt text: Two panels. Panel A is a bar chart of distance-axis combined standardized statistics for the five redox-node genes and CDKN1A, each against its abundance-matched null drawn as a horizontal line, with open markers showing the capture-area clustered value; MTARC1 falls below its clustered null while CDKN1A stays above both. Panel B is a grouped bar chart of standardized carriage residuals for the same genes under the registered and amended models, with dashed lines at the shortlist entry level of 1.8233 that none of the redox genes reaches and CDKN1A reaches under both models.

Table 1. Negative, self-correcting and refuted outcomes. Each registered or expected result the data did not deliver, with the values that settle it and its source. Related to Data S3.

outcome	what was registered or expected	what the data give	source file
Two registered confirmatory predictions refuted	Z-A family: flatness along the lobular axis for six genes	Four replicate it. POR +3.4151 against a null 97.5th percentile of 2.0861 and SCD1 +11.0210 against 2.4251 exceed the threshold they were registered to fall below; the register marks both as discrepancies	outputs/ZA_D3_verdict.tsv
Most components never analyzed	All tissue components of cohort D3 that clear the gates	Of 69 tissue components in cohort D3, 36 fell below the 200-spot minimum and 11 failed the P4 landmark gate, leaving the 22 that enter the lobular analysis. P4 failed components of 838 and 641 spots, so it is not only a size filter	outputs/d3_rerun/D3_components.tsv
Distance axis not analyzable as registered	100 spots at distance zero and 100 beyond distance three	Exactly 1 of 22 components met it, and it failed F2c, so 0 of 22 passed as registered. Amendment A-F1 lowered the requirement to 30 spots and rebinned the axis. It did not relax F2c, which excludes 7 of 22 components, 4 of them among the 14 that clear the amended size gate	Table1_component_gates.tsv (Data S4)
Independent replication stopped at the premise check	Three of five D1 sections to pass before any replication claim	Two passed and three failed: two on size, one of those also on F2c, and JBO018 on F2c and D3b. In JBO018 the hepatocyte identity score rose toward the putative scar at -4.945 with a quintile ratio of 1.94 against a requirement of 2.0, so that section most likely has no scar. The two passing sections were run and their verdicts are in Data S3, on two units, below the registered requirement	outputs/d1_replication/D1_gate.tsv
Post hoc explanation registered and tested	Hypothesis Z3: pericentral scar initiation cancels the parenchymal density effect for pericentral genes	The registered primary endpoint is inconclusive at -1.6599, with a detection limit of 0.585, while two registered variants point toward it, a reduced zonation index at -2.7703 and an ordered-class test at a permutation p of 0.0016. Neither supported nor refuted	outputs/z3/Z3_verdict.tsv
Registered exploratory family omitted from an earlier draft	Senescence family F-4: CDKN1A, CDKN2A, GDF15, SERPINE1 and a module	CDKN1A exceeds its null; CDKN2A not analyzed, one component supplying a matched null; GDF15 and SERPINE1 are flat at -0.8216 and +0.5065, with detection limits of 0.994 and 0.752, and the module is inconclusive at +1.7325 against 1.9405, with a detection limit of 0.868	Table2_distance_endpoints.tsv (Data S2)
Registered structural prediction wrong	A large residual shortlist on the lobular axis and a small one on the distance axis	The observed shortlists are 65 and 488, the reverse. Shortlist size reflects how much structure the model removed, not how much signal it left	outputs/carriage/CC_C3_residuals_*.tsv
Registered gates never run on the analyzed cohort	Gates D3a and D3b, testing whether a component's top collagen quintile is scar at all	Both were written for the replication cohort D1 and were never run on the D3 components; D3b is the gate that caught JBO018 at 1.94, so the mask of this paper's primary axis rests on the fibrosis score alone	PREREG_D1_replication, outputs/d1_replication/

Registered model incomplete on both axes	A two-term compartment model with no endothelial term	Sinusoidal endothelial markers are enriched 77.9-fold in the registered lobular shortlist and 11.4-fold in the registered distance shortlist; adding an endothelial term reduces these only to 12.8-fold and 8.4-fold	outputs/carriage/, Supplementary Results

Table 2. The estimator-by-reference combinations against the registered premise check. The check, fixed before any run, requires a pooled median hepatocyte fraction of at least 0.40 with hepatocytes the most abundant class. Every row that ran used the same 9,714 spots, the same 22 components and the same deposited reference exports; the refused row returned no fractions at all. The last three rows are the amended variant that drops the classes built on fewer than 25 nuclei, added under amendment C-D4. "Median" is the pooled median hepatocyte fraction and "passing" the number of components clearing the check. "Next class" and "immune", added under addendum C-D18, are not part of the check: the first gives each cell's leading non-hepatocyte class with its pooled median, the second the pooled median summed over all immune classes, so that the composition the check passed is visible. Among the four cells that clear it those totals are 0.0044, 0.3118, 0.0936 and 0.0032, and the leading non-hepatocyte class is a different one in each. Every class of every cell is in Data S5. RCTD was added under amendment C-D14 and the CARD row on the amended arm under C-D16. The albumin reading behind the unmodified single-nucleus arm's failure, 11,813.3 counts per million in resident NK nuclei, is twenty reads. Related to Figure 2 and Data S5.

estimator	reference	median	most abundant	next class	immune	passing	verdict
per-spot non-negative least squares	dissociated compilation, 17 classes	0.5639	hepatocytes	cholangiocytes (0.1906)	0.0044	22 of 22	passes
RCTD 2.2.1	dissociated compilation, 17 classes	0.4445	hepatocytes	plasma cells (0.1491)	0.3118	18 of 22	passes
CARD 1.1	dissociated compilation, 17	0.3137	fibroblasts, at 0.3937	fibroblasts (0.3937)	0.0862	0 of 22	fails

	classes						
per-spot non-negative least squares	single-nucleus subset, 10 classes	0.0000	resident NK	resident NK (0.4647)	0.9456	0 of 22	fails
CARD 1.1	single-nucleus subset, 10 classes	0.0950	resident NK	resident NK (0.3829)	0.7770	0 of 22	fails
RCTD 2.2.1	single-nucleus subset, 10 classes	not analyzed	not analyzed	not analyzed	not analyzed	not analyzed	not analyzed: the estimator refuses this reference, five of its ten classes holding fewer cells than its own minimum of 25, at 11, 2, 2, 2 and 1, so no fractions are returned to check. The minimum was not lowered, because overriding an estimator's default to make it match another estimator is tuning
per-spot non-negative least squares	single-nucleus subset, classes under 25 cells dropped (amendment C-D4)	0.2357	fibroblasts	fibroblasts (0.7078)	0.0072	0 of 22	fails
RCTD 2.2.1	single-nucleus subset, classes under 25 cells dropped (amendment C-D4)	0.7526	hepatocytes	macrophages (0.0936)	0.0936	22 of 22	passes
CARD 1.1	single-nucleus subset, classes under 25 cells dropped (amendment C-D4)	0.5871	hepatocytes	fibroblasts (0.3640)	0.0032	22 of 22	passes

Across the eight least-squares variants, four profile normalizations crossed with two response scalings, the pooled median lies between 0.4669 and 0.6130 on the dissociated compilation, is 0.0000 on the unmodified single-nucleus subset, and lies between 0.1541 and 0.2357 with fibroblasts most abundant on the amended arm. The two response scalings are redundant, so these are four distinct values per arm, not eight. No verdict changes under any variant, so least squares fails the amended arm for reasons that are not a choice of normalization (amendments C-D16, C-D17). Under C-D17 the sweep reports the registered pooled median over spots, and its unit-normalized panel-CPM cell reproduces the primary run exactly; the earlier figures of 0.4763 to 0.6202 and 0.4640 to 0.6104 were medians of per-component medians and are superseded. Across premise thresholds from 0.25 to 0.55 one cell changes verdict, RCTD on the dissociated compilation, whose 0.4445 sits just above the registered 0.40 and falls below 0.45. Every other cell holds at every threshold in that range.

Table 3. Six limits on reading the carriage contrast as compositional. Each limit is stated with the quantity that measures it. None is removed by the data in hand. Related to Figure 3 and Figure 4.

limit	what it is	the quantity that measures it
the components were selected on a quantity close to the endpoint	gate F2c admits a distance-axis component only if a hepatocyte identity score rises with distance from scar and beats a matched control, and the primary endpoint is in substance the gene-level version of that gate; the lobular axis has no gate on the same quantity, although it has a landmark-orientation gate that 11 of 33 sufficiently sized components failed	Spearman correlation between the combined statistic and hepatocyte carriage is +0.378 on the distance axis against +0.031 on the lobular axis; under amendment C-D13 the twelve refused components still give 0.11006 against the gated 0.15452, which is 0.712 of it, so the selection attenuates the contrast rather than creating it
the permutation null does not identify the predictor as composition	it breaks the gene-to-carriage link while leaving the expression covariate and the trend distribution intact, so clearing it shows only that carriage is not noise	substituting log reference counts, which carry no compositional information,

		gives 0.03419 against 0.00004, 24 percent of the headline difference, at an axis ratio of 855 against 16.5
the denominator is rivaled by a nuisance the model omits	the number of components in which a gene was detected is not in the registered model	it explains 0.00579 of the lobular rank variance beyond the expression term, 0.62 times the lobular increment; carrying it leaves the increments at 0.15401 and 0.00934
the axis ratio is not identified under component sampling	recombining the lobular axis over 500 random subsets of ten of its twenty-two components, the size of the distance axis, moves the ratio over more than an order of magnitude, while leaving its sign untouched	under amendment C-D11 the ratio runs from 5.19 to 109.08 for 95 percent of draws against an observed 16.57, which sits at the 48th percentile; the smallest of the 500 draws is 3.87
the ratio depends on how the predictor is parameterized	carriage is the prior art's cell-type specificity reweighted by the class transcript totals of the reference compilation, and the two carry the same information about a gene	under amendment C-D12 specificity gives 0.22437 and 0.02655 against carriage's 0.15448 and 0.00934 in the same model, an axis ratio of 8.45 against 16.55
the increment is on the rank scale, and the interval covers gene sampling only	what is explained is the variance of a trend statistic's rank across genes, not the magnitude of any gradient in tissue; the bootstrap treats as exchangeable genes that share both a compartment and a trend	recomputed on the raw combined deviate the values are 0.12983 and 0.0024, a ratio of 54

Table 3 note. The carriage increment carries a 2,000-draw bootstrap over genes and a 1,000-permutation null. The matched contrast rests on the ten distance components clearing gate F2c, which requires hepatocyte identity to rise with distance, and the amended size gate. On the twelve refused components, seven failing F2c and five refused on size alone, the increment is 0.11006 against the gated 0.15452. Cell-type specificity, the prior art carriage reweights, does better in the same model under C-D12 and halves the ratio, so carriage is one parameterization of it. Shuffling carriage across genes collapses the increment to a 97.5th percentile of 0.00062 and 0.00069; a stricter null permuted within abundance strata, under C-D10, leaves it 55 times its null of 0.0028 rather than 211 times 0.00073. The matched comparison is built on the ten distance components that clear F2c with the amended size gate, of 15 clearing F2c and 14 clearing size. Resampling under C-D9 gives shares from 0.4483 to 0.7020 for 95% of draws, the gated ten at the tenth percentile.

So the implied ratio runs from 1.06 to 1.65 for 95% of draws, with the observed 1.50 at their 90th percentile. The schema needs only an annotated reference, a coordinate and a trend statistic. The implementation names liver at six points, each of which a user in another tissue would respecify: CYP2E1 as the pericentral anchor of the lobular axis; a 22-gene fibrillar collagen score as the fibrosis score that defines the distance axis; the hepatocyte identity panel behind gate F2c; the hepatocyte fraction of at least 0.40 in the premise check; the 16-gene hepatocyte marker panel of the marker baseline; and the 127-marker lineage panel used by the least-squares fit.

Table 4. One endpoint that dissolves and one that holds, under two restrictions. Each row gives one endpoint as registered and under each restriction. "Z" is the combined standardized statistic, "null" its abundance-matched null at the 97.5th percentile, and "limit" the per-unit detection limit. Where the two matching rules disagree, the fibrotic-distance registration designates the fixed-tolerance result as governing. Related to Figure 5 and Data S2.

gene	analysis	Z	null	limit	verdict
MTARC1	as registered, quantile rule	+3.129	2.1406	0.677	away from scar
MTARC1	as registered, governing fixed rule	+3.2813	2.54	0.803	away from scar
MTARC1	capture areas as units	+1.7816	1.9181	0.783	inconclusive
MTARC1	documented areas and clustering together	+1.9115	1.9208	0.96	inconclusive
MTARC2	as registered, quantile rule	+2.2013	1.7762	0.562	away from scar
MTARC2	capture areas as units	+1.3747	1.5364	0.627	inconclusive
CYB5R3	as registered, quantile rule	-2.1885	2.2045	0.697	inconclusive
CYB5R3	documented areas only	-1.1317	2.2071	0.834	flat
CDKN1A	as registered, quantile rule	+3.2084	2.1607	0.683	away from scar
CDKN1A	documented areas and clustering together	+2.0724	1.9538	0.977	away from scar

Table 5. The procedure run unchanged on a non-liver section, under transfer registration T1. Sample 151673 of a public human dorsolateral prefrontal cortex Visium

series [19], 3,639 spots. The statistics, binning, mask quantile, size gate and matched-null rules are imported from the liver run's module. Four slots name the tissue: the mask panel, the identity panel behind the F2c analog, the endpoint set and the exclusion list for null draws. Under addendum C-D19 every cortex endpoint that did not clear is reported as inconclusive rather than flat: the registered verdict rule refutes only below the 84th percentile of the matched null, and the T1 run deposited the 2.5th and 97.5th percentiles only, so the refutation limb is not evaluable here and the rule's own default applies. The transfer therefore demonstrates the axis, the gates and the matched-null comparison, not the three-valued verdict rule. "Deviate" is the Jonckheere-Terpstra statistic across the five distance bins, "null" the bound of its abundance-matched null on the side the deviate falls, the 97.5th percentile for a positive deviate and the 2.5th for a negative one, under the governing fixed-tolerance rule, and "limit" the detection limit for that endpoint. Related to Data S1.

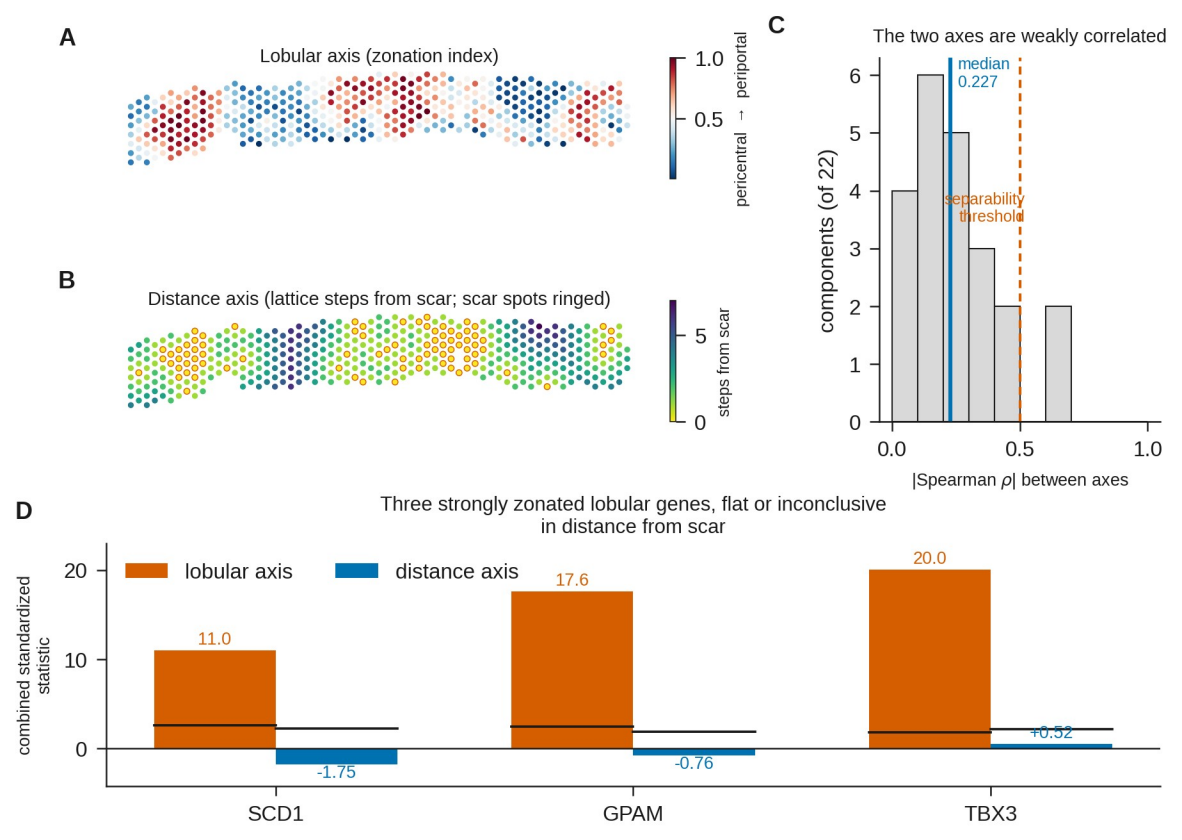
gate or endpoint	expected	deviate	null	limit	outcome
masked fraction	0.20 by construction				0.2001, F2a vacuous as in liver
first and last bin sizes	at least 30 each				728 and 2,561, passed
F2b, mask panel vs distance	negative	-48.278			passed
F2c, identity vs distance	positive, above its matched control	28.52	13.277		passed
CUX2	away from white matter	+8.1697	6.4132	3.6573	away from white matter
LAMP5	away from white matter	+6.8518	6.3162	3.4053	away from white matter
RASGRF2	away from white	+8.5491	9.658	4.536	inconclusive

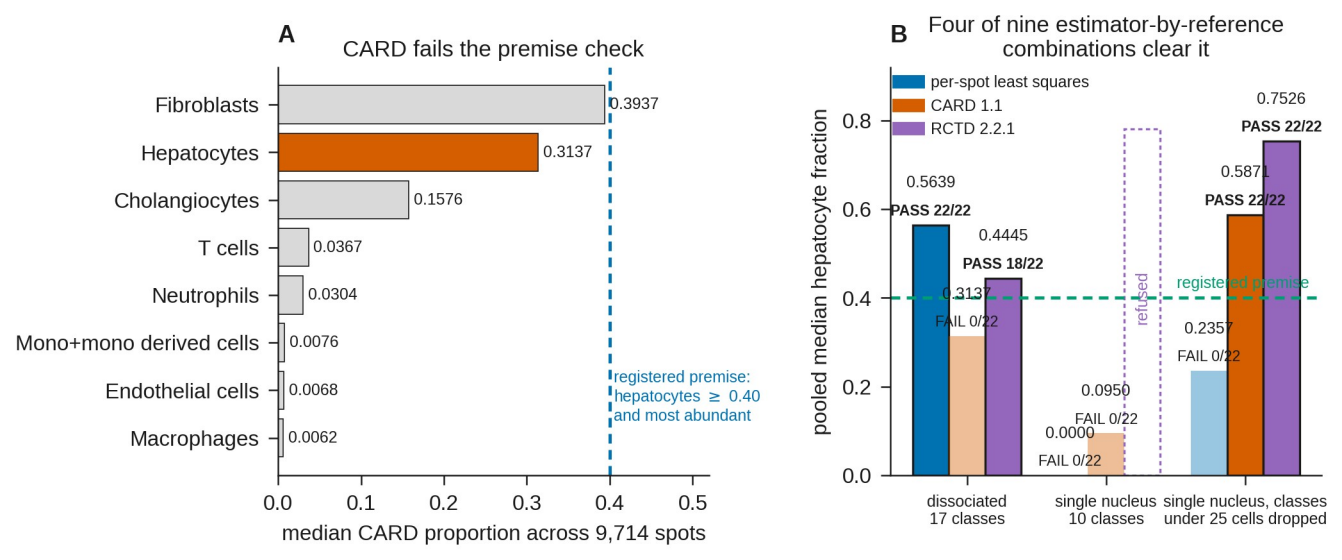
	matter				
CARTPT	away from white matter	+9.1276	9.3629	4.6783	inconclusive
SEMA3E	toward white matter	-3.0798	-2.9943	1.9179	toward white matter
TLE4	toward white matter	-5.4393	-5.8002	2.8441	inconclusive
FEZF2	toward white matter	+0.9592	2.5687	1.6671	inconclusive
NTNG2	toward white matter	+0.7482	5.3518	3.0995	inconclusive
PCP4	toward white matter	+20.5345	10 matched controls only		not analyzed under the governing rule, and contrary to expectation under the other
ALB	flat, being liver				not testable below the non-zero rate floor

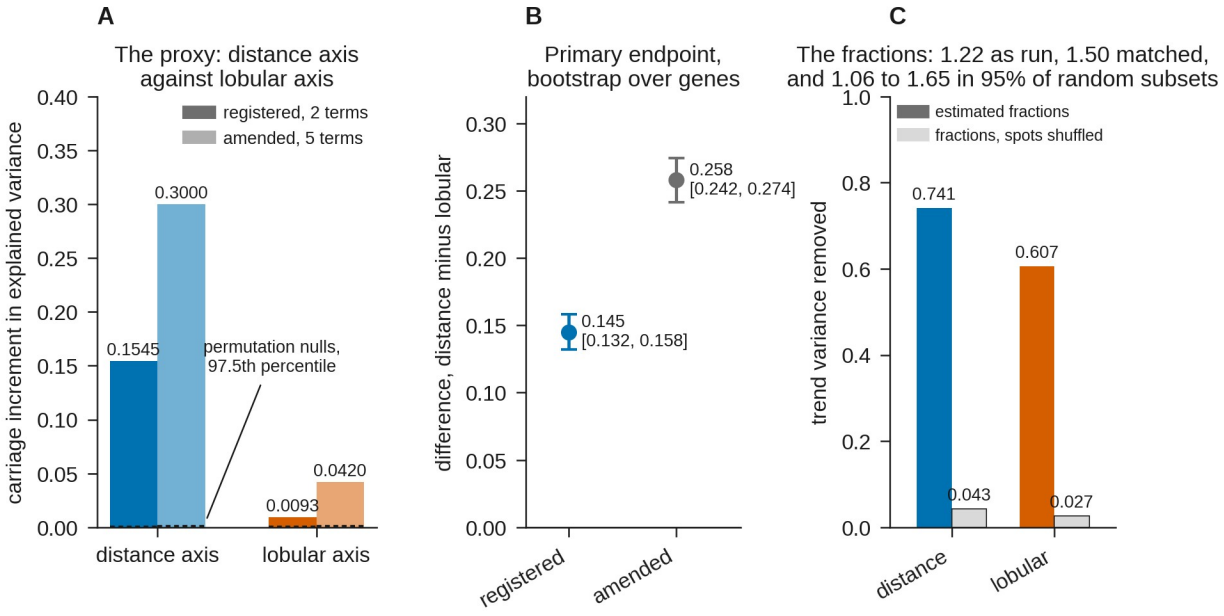
Table 6. What to check before reading a spatial trend as biology, with what each check returned here. The schema is tissue-agnostic; the values in the third column are this cohort's. A check that fails is a reason to stop, not a parameter to tune.

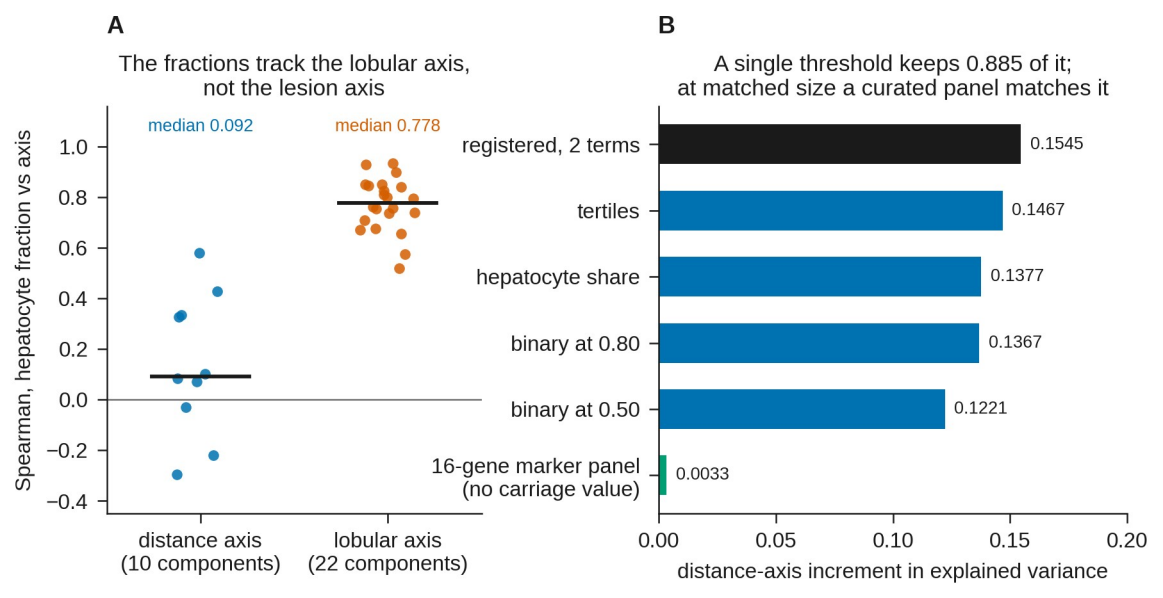
check	what it asks	here	if it fails
the reference premise, first limb	before estimated fractions enter a model as covariates, does the reference recover the tissue's dominant class, by a median fraction fixed in advance and by being the most abundant class	four of nine estimator-by-reference combinations cleared a median of 0.40 with hepatocytes most abundant, and one estimator refused a reference outright	report the combination as not analyzed; do not lower the threshold, which rescues no failing combination here between 0.25 and 0.55
the reference premise, second limb	do the non-dominant classes fall in a range the tissue allows, which the first limb never inspects	the four combinations that cleared the first limb put their leading non-hepatocyte class at cholangiocytes 0.1906, plasma cells 0.1491, macrophages 0.0936 and fibroblasts 0.3640, with immune totals from 0.0032 to 0.3118, so they are not describing one tissue and the first limb cannot see it (C-D18)	state the full composition, not the dominant class alone, and treat a cell that clears the first limb as licensed for the dominant class only; any model adjusting on the other columns owes this limb
coordinate and composition collinearity	is the spatial coordinate itself tracking the estimated fractions	the median correlation is 0.7775 on the lobular axis against 0.0925 on the	do not adjust on the collinear axis; residualizing there removes the trend and

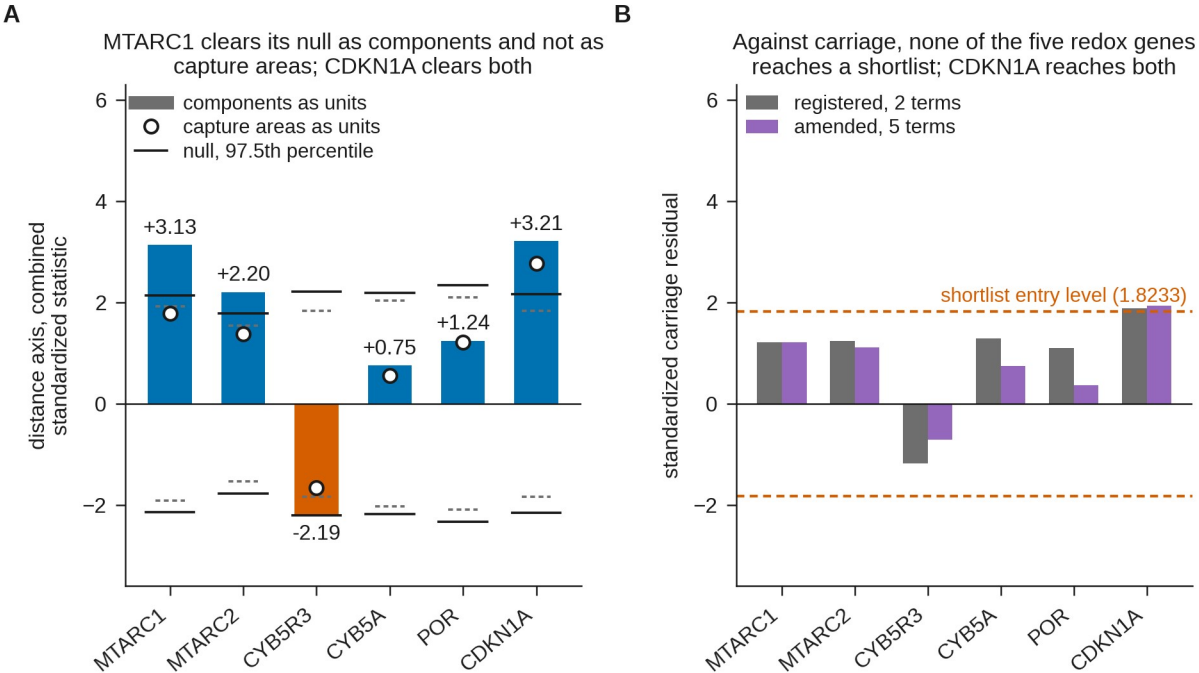
		lesion-anchored one	leaves a reversed remainder
what the null controls	does the empirical null control which cells express the gene, or only how much it is expressed	an abundance-matched null passed MTARC1, which neither of two further checks then supported, while CDKN1A survived both	add a reference-specificity residual and a capture-area clustered null before reporting
the unit of replication	are the units independent, or are several cut from one capture area	the ten distance-axis components come from six capture areas, three of them from one	cluster the null within capture area so the null matches the estimator, and report the clustered verdict
whether a magnitude is identified	does the size of the contrast survive resampling the units it was measured on	neither route's magnitude does. Over 500 random ten-component subsets the proxy's axis ratio runs from 5.2 to 109.1 for 95% of draws, and the fractions' from 1.06 to 1.65 for 95% of draws with 2 of 500 falling to 0.997, so the direction holds in all but two draws and no size is pinned down	report the direction and the interval, not the point estimate


A
Lobular axis (zonation index)
1.0
0.5
pericentral → periportal
B
Distance axis (lattice steps from scar; scar spots ringed)
5
0
steps from scar
C
The two axes are weakly correlated
median
0.227
separability
threshold
components (of 22)
0
1
2
3
4
5
6
0.0
0.5
1.0
|Spearman ρ| between axes
D
Three strongly zonated lobular genes, flat or inconclusive
in distance from scar
lobular axis
distance axis
combined standardized
statistic
20
10
0
11.0
-1.75
17.6
-0.76
20.0
+0.52
SCD1
GPAM
TBX3


A
CARD fails the premise check
Fibroblasts
Hepatocytes
Cholangiocytes
T cells
Neutrophils
Mono+mono derived cells
Endothelial cells
Macrophages
0.3937
0.3137
0.1576
0.0367
0.0304
0.0076
0.0068
0.0062
registered premise:
hepatocytes ≥ 0.40
and most abundant
0.0
0.1
0.2
0.3
0.4
0.5
median CARD proportion across 9,714 spots
B
Four of nine estimator-by-reference
combinations clear it
per-spot least squares
CARD 1.1
RCTD 2.2.1
pooled median hepatocyte fraction
0.0
0.2
0.4
0.6
0.8
0.5639
PASS 22/22
0.4445
PASS 18/22
0.3137
FAIL 0/22
0.0950
FAIL 0/22
0.0000
FAIL 0/22
refused
registered premise
0.2357
FAIL 0/22
0.5871
PASS 22/22
0.7526
PASS 22/22
dissociated
17 classes
single nucleus
10 classes
single nucleus, classes
under 25 cells dropped


A
The proxy: distance axis against lobular axis
registered, 2 terms
amended, 5 terms
carriage increment in explained variance
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.3000
0.1545
permutation nulls, 97.5th percentile
0.0420
0.0093
distance axis
lobular axis
B
Primary endpoint, bootstrap over genes
difference, distance minus lobular
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.258
[0.242, 0.274]
0.145
[0.132, 0.158]
registered
amended
C
The fractions: 1.22 as run, 1.50 matched, and 1.06 to 1.65 in 95% of random subsets
estimated fractions
fractions, spots shuffled
trend variance removed
1.0
0.8
0.6
0.4
0.2
0.0
0.741
0.043
0.607
0.027
distance
lobular


A
The fractions track the lobular axis,
not the lesion axis
median 0.092
median 0.778
Spearman, hepatocyte fraction vs axis
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
distance axis
(10 components)
lobular axis
(22 components)
B
A single threshold keeps 0.885 of it;
at matched size a curated panel matches it
registered, 2 terms
0.1545
tertiles
0.1467
hepatocyte share
0.1377
binary at 0.80
0.1367
binary at 0.50
0.1221
16-gene marker panel
(no carriage value)
0.0033
0.00
0.05
0.10
0.15
0.20
distance-axis increment in explained variance


A
MTARC1 clears its null as components and not as capture areas; CDKN1A clears both
components as units
capture areas as units
null, 97.5th percentile
distance axis, combined standardized statistic
+3.13
+2.20
-2.19
+0.75
+1.24
+3.21
MTARC1
MTARC2
CYB5R3
CYB5A
POR
CDKN1A
B
Against carriage, none of the five redox genes reaches a shortlist; CDKN1A reaches both
registered, 2 terms
amended, 5 terms
standardized carriage residual
shortlist entry level (1.8233)
MTARC1
MTARC2
CYB5R3
CYB5A
POR
CDKN1A

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigureS1.tiff
- Supplementarymaterial.docx
- DataS2toS5.pdf