Biometrics 000, 1–22 DOI: 000
000 0000

A Zero-Inflated Hierarchical Generalized Transformation Model to Address Non-Normality in Spatially-Informed Cell-Type Deconvolution

Hunter J. Melton[1], Jonathan R. Bradley[2,], and Chong Wu[3,***]**

[1]Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire, U.S.A.

[2]Department of Statistics, The University of Missouri, Columbia, Missouri, U.S.A.

[3]Department of Biostatistics, University of Texas MD Anderson Cancer Center, Houston, Texas, U.S.A.

**email:* bradleyjr@missouri.edu

***email:* CWu18@mdanderson.org

Summary: Oral squamous cell carcinomas (OSCC), the predominant head and neck cancer, pose significant challenges due to late-stage diagnoses and low five-year survival rates. Spatial transcriptomics offers a promising avenue to decipher the genetic intricacies of OSCC tumor microenvironments. In spatial transcriptomics, Cell-type deconvolution is a crucial inferential goal; however, current methods fail to consider the high zero-inflation present in OSCC data. To address this, we develop a novel zero-inflated version of the hierarchical generalized transformation model (ZI-HGT) and apply it to the Conditional AutoRegressive Deconvolution (CARD) for cell-type deconvolution. The ZI-HGT serves as an auxiliary Bayesian technique for CARD, reconciling the highly zero-inflated OSCC spatial transcriptomics data with CARD's normality assumption. The combined ZI-HGT + CARD framework achieves enhanced cell-type deconvolution accuracy and quantifies uncertainty in the estimated cell-type proportions. We demonstrate the superior performance through simulations and analysis of the OSCC data. Furthermore, our approach enables the determination of the locations of the diverse fibroblast population in the tumor microenvironment, critical for understanding tumor growth and immunosuppression in OSCC.

Key words: Bayesian, spatial transcriptomics, uncertainty quantification, zero-inflation.

This paper has been submitted for consideration for publication in *Biometrics*

1 Introduction

Oral squamous cell carcinomas (OSCC) are the most common head and neck cancer, comprising over 90% of oral cavity mucosal malignancies (Markopoulos, 2012). OSCC progression is often undetected until advanced stages, contributing to five-year survival rates of approximately 50% (Markopoulos, 2012). Spatial transcriptomics (ST) technologies, which provide spatial localization of gene expression (Moses and Pachter, 2022), offer new opportunities to characterize tumor cellular and molecular landscapes. The recent OSCC ST dataset by (Arora et al., 2023) enables detailed investigation of the tumor microenvironment (TME) and spatial distribution of cell types, providing novel perspectives on tumorigenesis and cellular interactions (Moses and Pachter, 2022).

However, the application of ST in OSCC research encounters notable hurdles with spatial barcoding-based methods, such as the 10X Visium platform used to capture the OSCC ST data (Arora et al., 2023). These methods collect gene expression measurements from tissue locations containing multiple cells of diverse types (Ståhl et al., 2016), effectively quantifying average expression levels across cells at each location. Consequently, cell-type deconvolution, which probabilistically infers cell type proportions at each location, is crucial for determining spatial cell type distributions and enabling subsequent analyses (Ma and Zhou, 2022).

Despite these advances, existing approaches often overlook zero inflation endemic to ST data (Li et al., 2021; Zhao et al., 2022). For example, CARD (Ma and Zhou, 2022) assumes the spatially resolved gene expression data are normally distributed, disregarding zero-inflation and ties, which may result in reduced performance due to mismatched data and modeling assumptions. While Zhao et al. (2022) argue that zero-inflation can be addressed by controlling for cell-type composition or by analyzing spatially homogeneous tissue with standard count models such as a Poisson or negative binomial, their analysis focuses on scenarios where spatial heterogeneity is controlled rather than modeled. In contrast, methods

like CARD that assume normality of the raw or minimally processed data may benefit from transformations that address zero-inflation and ties simultaneously with spatial modeling.

Zero-inflation is substantial in the 12 OSCC samples (Arora et al., 2023), where sparsity ranges from 86%-91%, a pattern common across ST datasets (Moncada et al., 2020). While the high sparsity could be directly addressed by a zero-inflated Poisson model (Lambert, 1992) or zero-inflated Negative Binomial model (Greene, 1994), these are challenging and time-consuming given the spatial context and data size—over 15 million points per sample. Deterministic transformations, such as $\log(1+\epsilon+x)$, may be applied to ST data prior to analysis, though their impact is minimal (Web Figure 3). Consequently, few spatial transcriptomic methods address zero-inflation (Li et al., 2021), and no cell-type deconvolution methods do so directly. Moreover, uncertainty quantification (UQ) in cell-type deconvolution often relies on model-based assumptions that may not align with the true distribution of the data. Fully Bayesian models enable posterior credible intervals for UQ but require computationally intensive MCMC methods that renders them impractical for high-dimensional ST datasets.

In classical statistics, when data diverge from model assumptions (often normality), deterministic transformations, such as Box-Cox (Box and Cox, 1964) or Yeo-Johnson power transformations, (Yeo and Johnson, 2000) are employed to align data with the assumptions pre-analysis. However, classical deterministic transformations struggle with data containing many ties , as ties remain after transformation, resulting in persistent non-normality. Contrastingly, the hierarchical generalized transformation (HGT) (Bradley et al., 2020; Bradley, 2022; Bradley et al., 2023) introduces a probabilistic, rather than deterministic, transformation to data processing. The HGT is a new class of noisy transformations that breaks ties by adding a small amount of noise to the data, which is modeled via a Bayesian transformation. The HGT, however, has yet to be developed for zero-inflated data. Here, we introduce the Zero-Inflated Hierarchical Generalized Transformation (ZI-HGT), a novel adaptation

designed to address zero inflation and ties in ST datasets. By integrating a type of zero-inflated Poisson model into the HGT framework, the ZI-HGT generates posterior replicates from the noisy transformation, mitigating zeros and ties to ensure transformed data better aligns with model assumptions (i.e., normality in CARD). The posterior replicates of the transformation are then used in place of the original data. The combination of zero-inflated modeling with HGT and CARD produces a Bayesian hierarchical model that enables UQ and is implementable on high-dimensional (over 15 million data points), multivariate spatial datasets, a scale at which current fully Bayesian alternatives for zero-inflated multivariate spatial data become computationally intractable (Musgrove et al., 2019).

The rest of the manuscript is organized as follows: Section 2 discusses the motivating OSCC data. Preliminary concepts and methodology are discussed in Sections 3 and 4, respectively. An empirical simulation study is presented in Section 5, which elucidates the viability of the model. Section 6 applies the model to OSCC data, presenting findings, and Section 7 provides conclusions and discussions. Figure 1 contains a schematic overview of the method.

2 Data and Motivation

We analyze oral squamous cell carcinoma spatial transcriptomic (ST) data (Arora et al., 2023), consisting of 12 samples surgically resected from 10 patients' oral cavities. Nine tissue layers were annotated by a pathologist: squamous cell carcinoma (SCC), lymphocyte-positive stroma, lymphocyte-negative stroma, normal mucosa, glandular stroma, muscle, keratin, and artery/vein, and an artifact layer. ST was conducted using the 10x Genomics Visium platform, which samples approximately 10 cells per location. Across all samples, reads were obtained for an average of $16,763$ unique transcripts over $2,073$ spatial locations. Additionally, our analysis incorporates a reference single-cell RNA-sequencing (scRNA-seq) dataset (GSE103322) (Puram et al., 2017). Data collection and preparation are detailed in Web Appendix A. As an eaxmple, we consider OSCC ST Sample 1 and note that the other

samples are similar (Web Table 1). Sample 1 contains read counts for 15,844 transcripts across 1,131 spatial locations. Web Table 2 presents gene expression read count frequencies (see Web Tables 3-13 for other samples). We highlight two key features of the dataset. First, zero inflation is extremely high: 16,384,736 of 17,919,564 (91.4%) read counts are zero. Second, the count data contains many ties. These characteristics underscore the limitations of current analytical approaches like CARD, which assumes normally distributed ST data, an assumption rendered untenable by zero-inflation and abundant ties. To address these challenges, we introduce a novel zero-inflated HGT (a noisy transformation aimed at breaking ties and addressing zero-inflation) in the following section to address the non-normality of the data before conducting cell-type deconvolution with CARD.

3 Preliminaries

3.1 Notation

Let $\mathbf{X}$ be the $G \times N$ gene expression matrix from an OSCC sample (Arora et al., 2023), containing transcript reads for G cell-type informative genes measured at N spatial locations. Denote the (i, j)-th element $X_{i,j}$. Let $\mathbf{H}(\mathbf{X})$, or $\mathbf{H}$ for brevity, be the noisy transformation of $\mathbf{X}$, $\mathbf{X}^{[0]}$ be a set of indicators for positive gene expression in $\mathbf{X}$ (i.e., $X_{i,j}^{[0]} = 1$ if $X_{i,j} > 0$), and $\mathbf{X}^{(B)}$ be a data augmentation variable that facilitates the implementation of a zero-inflated model without expensive Markov chain Monte Carlo (MCMC) sampling. We employ corresponding scRNA-seq data (Puram et al., 2017) as a deconvolution reference and select informative genes for distinguishing between cell types. Let $\mathbf{B}$ be the $G \times K$ single-cell gene expression matrix (the reference basis matrix), where G is again the number of cell-type informative genes and K is the number of cell types. Lastly, let $\mathbf{V}$ be the $N \times K$ cell-type proportions matrix, where the rows of $\mathbf{V}$ contain the proportions of the K cell types for each spatial location. In the OSCC ST data, G and N may vary slightly from sample-to-sample, but $K = 14$. Futhermore, we introduce the following shorthand notation

for probability distributions. "Point Mass" is a short-hand for the point mass distribution, and "TruncPoisson" is a short-hand for the zero-truncated Poisson distribution.

3.2 Review of CARD

CARD (Ma and Zhou, 2022) is an empirical Bayesian model that applies the following non-negative matrix factorization model for $\mathbf{X}$:

$$\mathbf{X} = \mathbf{B}\mathbf{V}' + \mathbf{E}, \tag{1}$$

where $\mathbf{E}$ is the $G \times N$ residual error matrix and $E_{i,j} \sim N(0, \sigma_e^2)$. B contains the average gene expression level in each cell type from the scRNA-seq reference data for all "cell-type informative" genes, classified as such if their mean expression level in any cell type is at least 1.25 log-fold higher than the average across all other cell types (see Web Appendix H for details). To consider spatial colocalization of cell types, a conditional autoregressive (CAR) (Besag, 1974) model serves as the prior for columns of the cell-type proportions matrix $\mathbf{V}$. For spatial location i and cell type k, the full-conditional distribution for $V_{i,k}$ is

$$V_{i,k} = b_k + \phi \sum_{j \neq i} W_{i,j}(V_{j,k} - b_k) + \epsilon_{i,k}; \quad i = 1, ..., N, \ k = 1, ..., K, \tag{2}$$

where $V_{i,k}$ denotes the proportion of cells at location i belonging to cell type k, b_k is the cell-type intercept representing the average cell-type composition for cell type k across all locations, and $W_{i,j}$ refers to the weight for learning the cell-type composition at location i using the information at location j. The parameter ϕ represents the spatial autocorrelation, determining the strength of the spatial correlation in the cell-type composition, and $\epsilon_{i,k}$ represents the residual error such that $\epsilon_{i,k} \sim N(0, \sigma_{i,k}^2)$. The CAR model represents the proportion of cells at location i in cell type k $V_{i,k}$ as a weighted sum of cell type k proportions at the other locations, enabling spatial sharing of cell-type composition information.

Notably, Equation (1) assumes normality of the) spatially resolved gene expression $\mathbf{X}$, which conflicts with the high sparsity and prevalent ties in the OSCC data, as highlighted in

Section 2. CARD incorporates location-level normalization, dividing each read count in X by the sum of the read counts at that location, in an attempt to address potential technical artifacts. While this procedure alleviates some ties, it does not impact zeros, so high sparsity (minimum 86% across the 12 OSCC samples) remains (Web Figure 43).

Rather than sample from the posterior predictive distribution (Equation (9) of the supplementary material of Ma and Zhou (2022)), CARD applies an algorithm that iterates through $\mathbf{V}_k$ and the hyperparameters to minimize the negative log-likelihood. This computationally efficient approach constructs point estimates of the cell-type proportions but precludes UQ.

3.3 Review of HGT

Models like CARD assume Gaussian data, which is unrealistic given the prevalence of ties in the OSCC dataset. Adding noise to the data produces a new (transformed) dataset without ties, rendering the Gaussian assumption more reasonable. To accomplish this, the HGT adds small amounts of noise to count data to break ties that fully accounts for added uncertainty from the noisy transformation. The HGT implements a fully Bayesian model that overfits the original data (Bradley, 2022), with posterior predictive replicates interpreted as the noisy transformation. By "overfitting," the posterior replicates are close (in absolute difference) on average to a deterministic transformation of the data on some scale. The motivation for overfitting is that the added noise should be "small" enough to break ties without distorting the underlying signal or introducing substantial bias into downstream analyses.

Recall that $\mathbf{H} = \mathbf{H}(\mathbf{X})$ denotes the noisy transformation of the original data $\mathbf{X}$, and $H_{i,j} = \{\mathbf{H}\}_{i,j}$. The current use of the HGT assumes $X_{i,j}\big|H_{i,j} \sim \text{Poisson}\,(H_{i,j})$ and $H_{i,j} \sim \text{Gamma}\,(\alpha, \kappa)$ so that the posterior distribution $H_{i,j}\big|X_{i,j} \sim \text{Gamma}\,(\alpha + X_{i,j}, \kappa + 1)$. When, for example, $\alpha \approx \kappa$, the posterior replicates of $H_{i,j}\big|X_{i,j}$ have mean $Z_{i,j}$ that overfits the original data $X_{i,j}$. This model "overfits" as the posterior summary recover the original data values. CARD is then applied to transformation $\mathbf{H}$ in place of $\mathbf{X}$. Consider the following

illustration. Let $\alpha = \kappa = 0.1$, and let $X_{i,j} = 1$. Then, as $H_{i,j}\big|X_{i,j} \sim \text{Gamma}\left(\alpha + X_{i,j}, \kappa + 1\right)$, $Z_{i,j} = E\left\{H_{i,j}\big|X_{i,j}\right\} = \frac{\alpha + X_{i,j}}{\kappa+1} = 1 = X_{i,j}$, so the transformation mean equals the original data. More generally, $Z_{i,j} = E\left\{H_{i,j}\big|X_{i,j}\right\} \to X_{i,j}$ as $\alpha, \kappa \to 0$ or $X_{i,j} \to \infty$.

In the HGT, the added noise is parameterized as presented previously rather than a simpler transformation for two key reasons. First, as discussed in Bradley et al. (2020), HGT leverages signal-to-noise cross-covariance for better prediction (see Web Appendix D for details). Second, the conjugacy of the transformation enables direct sampling of the transformation posterior, significantly reducing computational costs. In fact, (Bradley, 2022) and (Bradley et al., 2023) compared the HGT framework to traditional Bayesian hierarchical models in various non-Gaussian settings, finding minimal differences in prediction and estimation but massive computational gains with the HGT. However, the presented HGT is suboptimal, as it fails to address the known zero-inflation in the OSCC ST data. Simulations suggest that ignoring zero-inflation in an HGT yields minimal improvement (Web Figure 3).

4 Methodology

4.1 Zero-inflated hierarchical generalized transformation (ZI-HGT)

When applying the ZI-HGT to CARD (Ma and Zhou, 2022), we address two features of the OSCC data: high zero inflation and ties. We consider two observed datasets: the original OSCC ST data $\mathbf{X}$ and $\mathbf{X}^{[0]}$. The $G \times N$ matrix $\mathbf{X}^{[0]}$ has (i,j)-th element $X_{i,j}^{[0]}$ representing an indicator of positive gene expression for gene i at location j, i.e., $X_{i,j}^{[0]} = 1$ for positive gene expression and $X_{i,j}^{[0]} = 0$ otherwise. We introduce $X_{i,j}^{(B)}$, where the superscript "B" stands for Bernoulli and $X_{i,j}^{(B)}$ will facilitate the implementation of a Bayesian zero-inflated model with a posterior distribution that is easy to directly sample from. We assume the hierarchical model:

$$X_{i,j}\big|X_{i,j}^{[0]} = 1, H_{i,j} \sim \text{TruncPoisson}\left(H_{i,j}\right), \tag{3}$$

$$X_{i,j}\big|X_{i,j}^{[0]}=0, H_{i,j}, X_{i,j}^{(B)} \sim \text{Point Mass}\,(0)\,, \tag{4}$$

$$H_{i,j}\big|X_{i,j}^{[0]}=0, X_{i,j}^{(B)} \sim \text{Point Mass}\left(X_{i,j}^{(B)}\right), \tag{5}$$

$$H_{i,j}\big|X_{i,j}^{[0]}=1 \propto (1-\exp\left(-H_{i,j}\right))\,\text{Gamma}(\alpha_1,\alpha_1), \tag{6}$$

$$\left(\frac{\exp\left(X_{i,j}^{(B)}\right)}{1+\exp\left(X_{i,j}^{(B)}\right)}\right)\Bigg|\alpha_0,\kappa_0 \sim \text{Beta}(\alpha_0,\kappa_0), \tag{7}$$

$$X_{i,j}^{[0]}\big|X_{i,j}^{(B)} \sim \text{Bernoulli}\left(\frac{\exp\left(X_{i,j}^{(B)}\right)}{1+\exp\left(X_{i,j}^{(B)}\right)}\right);\quad i=1,...,N,\ j=1,...,G. \tag{8}$$

"$H_{i,j}\big|X_{i,j}^{[0]} = 1 \propto (1-\exp\left(-H_{i,j}\right))\,\text{Gamma}(\alpha_1,\alpha_1)$" represents the conditionally specified and unnormalized prior:

$$f(H_{i,j}\big|X_{i,j}^{[0]}=1) = (1-\exp\left(-H_{i,j}\right))\,\frac{\alpha_1^{\alpha_1}}{\Gamma(\alpha_1)}H_{ij}^{\alpha_1-1}\exp(-\alpha_1 H_{ij}), \tag{9}$$

where a gamma probability density function is multiplied by $(1-\exp\left(-H_{i,j}\right))$ to offset the normalizing constant in the truncated Poisson likelihood. One could similarly replace the truncated Poisson with the truncated negative binomial and modify the prior in (6) to be a weighted logit-beta prior distribution. The negative binomial is popular for modeling counts as it allows modeling over-dispersion. However, the HGT is a class of noisy transformations, where the transformations arise from a posterior distribution of a model that overfits the data, and posterior replicates are used as transformed data in another model (i.e., CARD). Questions of overdispersion are hence more appropriate for the model applied to the transformations (i.e., CARD), and CARD accounts for overdispersion via spatial covariances.

In Equation (3), non-zero count-valued OSCC ST data is modeled as truncated Poisson ($\geqslant 1$) distributed, and in Equation (4) zero-valued OSCC ST data is modeled as a point mass at zero. Zero-inflated models adopt a similar strategy but allow some zeros to be modeled in a Poisson mixture component. We instead use the truncated Poisson and unnormalized prior

for $H_{i,j}$, $(1-\exp(-H_{i,j}))\,\text{Gamma}(\alpha_1,\alpha_1)$ to achieve a closed-form posterior distribution for direct sampling (Equations (10) and (11) below). The prior distributions for the transformation $H_{i,j}$, Equations (6) and (7), are conditionally specified on the observed $X_{i,j}^{[0]}$. This choice produces a posterior distribution for $H_{i,j}$ with a closed form and the desired overfitting property (see Section 3.2 for motivation, and Web Figure 42 for a posterior transformation $\mathbf{H}$ of OSCC Sample 1 split by raw data value $X_{i,j}$).

The posterior transformation in closed-form (derived in Web Appendix C) is given by

$$H_{i,j}\big|X_{i,j}^{[0]}=1, X_{i,j}\sim \text{Gamma}(\alpha_1+X_{i,j},\alpha_1+1);\quad i=1,...,N,\ j=1,...,G, \tag{10}$$

$$\frac{\exp(H_{i,j})}{1+\exp(H_{i,j})}\Bigg|X_{i,j}^{[0]}=0, X_{i,j}\sim \text{Beta}(\alpha_0,\kappa_0+1);\quad i=1,...,N,\ j=1,...,G, \tag{11}$$

which has mean $\frac{\alpha_1+X_{i,j}}{\alpha_1+1}$ when $X_{i,j}^{[0]}=1$ and mean $\frac{\alpha_0}{\alpha_0+\kappa_0+1}$ when $X_{i,j}^{[0]}=X_{i,j}=0$, overfitting the data. The ZI-HGT can be sensitive to hyperparameter choices, so discrete ranges for α_0, κ_0, and α_1 are considered. Specific values are chosen to minimize the WAIC (Watanabe, 2010), a generalized version of the AIC for model selection. In simulations, WAIC-chosen hyperparameters result in similar performance to oracle-chosen (RMSE-minimizing) hyperparameters (Web Figure 2). Further details and practical recommendations for choosing hyperparameters are discussed in Web Appendix C.

Upon marginalizing $X_{i,j}^{[0]}$ in (4) and (3), we come to the expression

$$X_{i,j}\big|H_{i,j},X_{i,j}^{(B)}\sim \pi_{i,j}^{*}\text{TruncPoisson}(H_{i,j})+(1-\pi_{i,j}^{*})\text{Point Mass}(0), \tag{12}$$

where $\pi_{i,j}^{*}=f\left(X_{i,j}^{[0]}=1\big|H_{i,j},X_{i,j}^{(B)}\right)$ (detailed in Web Appendix E), and the expression in Equation (12) is of the form of a zero-inflated model (Lambert, 1992). While traditionally a Poisson distribution would replace the truncated Poisson, our goal differs from standard zero-inflated models. We aim to produce posterior replicates of *a transformation* for subsequent analysis rather than directly using posterior summaries of $H_{i,j}$. Using posterior replicates enables uncertainty quantification of estimated cell-type proportions downstream. Thus, the

truncated Poisson distribution is appropriate, yielding Equations (10) and (11) and hence a noisy transformation $H_{i,j}$ that approximates the original data $X_{i,j}$ on some scale.

4.2 Applying ZI-HGT to cell-type deconvolution

Denote the c-th replicate of $\mathbf{H} = \{H_{i,j}\}$ from Equations (10) and (11) with $\mathbf{H}^{[c]} = \{H_{i,j}^{[c]}\}$ for $c = 1, ..., C$. Following the HGT strategy, we assume

$$\mathbf{H}^{[c]} = \mathbf{B}\mathbf{V}' + \mathbf{E}, \quad c = 1, ..., C, \tag{13}$$

where $\mathbf{E} = \{E_{i,j}\}$ represents independent normal random variables with constant variance, $\mathbf{V}$ is the non-negative cell-type proportions with columns following a conditional autoregressive (CAR) (Besag, 1974) model, and $\mathbf{B}$ is the reference basis matrix. Algorithm 1 details the implementation.

The ZI-HGT model is defined by Equations (3-8), (13) and (2) (Bradley, 2022). From Bradley (2022), the BHM expression requires careful consideration of normalizing constants as Equations (3-8) condition on $\mathbf{H}^{[c]}$, while Equation (13) includes $\mathbf{H}$ but not as a condition. While the ZI-HGT alone does not model spatial variability, the combined ZI-HGT + CARD does through CARD. Web Appendix C provides the complete BHM for ZI-HGT + CARD.

Since no MCMC is required, the full method with CARD can be run efficiently by generating each of the posterior predictive replicates of the transformed data $H_{i,j}^{[c]}$ then computing the cell-type proportions estimates $\mathbf{V}^{[c]} = \widehat{\mathbf{V}}\left(\mathbf{H}^{[c]}\right)$ in parallel. The estimator $\widehat{V}$ is the function that minimizes the least squares $E\left\{||\widehat{\mathbf{V}}(\mathbf{H}) - S(\mathbf{H})||_2^2|\mathbf{X}\right\}$, where S is a generic real-valued estimator and $||\cdot||_2$ is the L_2 norm. The transformed OSCC data replicates $\mathbf{H}$ align significantly better with CARD's normality assumption than the original data $\mathbf{X}$ (Figure 2), containing no zeros or ties and substantially reducing the large right tail. While normality tests like Kolmogorov-Smirnov would be useful, they require independent data, whereas our data are multivariate and spatially dependent.

Consequently, ZI-HGT-generated estimates are more accurate at no additional computa-

tional cost, as demonstrated in our simulation study. We recommend $C = 100$ replicates. UQ is accomplished through replicates of Algorithm 1, the iterated variance formula, and a Taylor series argument to construct pointwise credible intervals for $\widehat{\mathbf{V}}$. UQ is detailed in Web Appendix F. Though it is not recommended due to the unclear cause of the zero-inflation, Web Appendix G presents the inverse transformation to return predictions back to the original scale of the data for completeness.

Algorithm 1: Applying the ZI-HGT to CARD

For $c \in \{1, ..., C\}$

[**1**] Sample posterior predictive transformations $H_{i,j}^{[c]}$ from $H_{i,j}|X_{i,j}^{[0]}, X_{i,j}$ via (11) and (10)

[**2**] Apply CARD to the transformed data: $V^{[c]} = \widehat{V}\left(H^{[c]}\right)$

[**3**] Obtain overall cell-type proportions estimate $\widehat{V} = \sum_{c=1}^{C} \widehat{V}^{[c]}/C$

5 Simulations

We conducted a rigorous simulation study using the OSCC data to evaluate the performance of the ZI-HGT + CARD across multiple settings, including varying sparsity levels and numbers of cell types in the ST data. We compared our approach with existing methods and a simple deterministic transformation to ensure comprehensive evaluation. The study focuses on OSCC Sample 2, which we consider representative of most samples, though sample choice has minimal impact (Web Figure 2). Sample 2 contains data from $1,749$ locations and $15,624$ transcripts. Web Appendix B provides additional details regarding the simulations.

5.1 Simulation Results

We first examined the impact of ST data sparsity on ZI-HGT + CARD performance, measured by the root mean square error (RMSE) in Figure 3. Higher sparsity in the simulated ST data resulted in stronger performance of the ZI-HGT + CARD relative to CARD alone. Sparsity levels of the data are determined the "library factor" and "ϕ factor" simulation hyperparameters; specific choices for both are presented in Figure 3 and detailed in Web Appendix B. At sparsity level 1 (median sparsity 89.8%, representing realistic ST data), the

median RMSE reduction of ZI-HGT + CARD beyond CARD was 6.6%. At sparsity level 2 (median sparsity 77.9%), the median reduction was 4.6%. This trend is expected and occurs as higher sparsity creates less normally distributed ST data, further deviating from CARD's assumptions and reducing its accuracy. The ZI-HGT addresses zero-inflation, so its impact increases as the ST data becomes "less normal." We discuss the performance of ZI-HGT + CARD primarily in terms of sparsity, but the median proportions of data points tied at 1, 2, 3, 4, and 5 are given for each sparsity level for additional context.

Next, we investigated how the number of distinct cell types affects performance (Web Figure 1). While median RMSE reduction remains consistent across varying cell type numbers, variance decreases as the number of cell types decreases. This expected outcome occurs because fewer cell types create more compact, simplified simulated ST data. We note that sparsity is unrelated to the number of included cell types, with median sparsity levels ranging from 88.2% to 89.2% across datasets with different cell type counts.

To ensure robustness given our focus on OSCC Sample 2, we examined the impact of sample choice for simulating ST data (Web Figure 2). We generated simulated ST data from SPARSim-generated scRNA-seq data using "realistic" settings (sparsity level 1). Our methods showed consistent RMSE reduction, with a median of 6.6%, though the variance was higher in Samples 8, 9, and 10, which were particularly homogeneous. These results reinforce our confidence in using OSCC Sample 2 as a representative choice for simulations.

We also compared ZI-HGT hyperparameters selected using WAIC vs. the oracle hyperparameters that minimize RMSE in simulated datasets based on each OSCC sample (Web Figure 2). Minimizing the WAIC yielded reasonable hyperparameter choices, with median RMSE reduction differences between WAIC-chosen and oracle models consistently small, nearly always within 1%. While the ZI-HGT can be somewhat sensitive to the choices of hyperparameters, hyperparameters that produce significantly worse performance consistently

had higher WAIC values (Web Figure 46), indicating that the WAIC selects hyperparameters well. Practical recommendations for choosing hyperparameters are discussed in Web Appendix C.

We compared the ZI-HGT + CARD to several alternatives: a simple deterministic transformation before CARD, HGT + CARD (the ZI-HGT without addressing zero-inflation), imputation with ALRA (Linderman et al., 2022) + CARD, denoising with MIST (Wang et al., 2022), and common cell-type deconvolution methods SPOTlight (Elosua-Bayes et al., 2021), SpatialDecon (Danaher et al., 2022), and STdeconvolve (Miller et al., 2022) (Web Figure 3). The deterministic transformation $log\,(1 + \epsilon + X_{i,j})$ is chosen to mimic the ZI-HGT: the log mitigates right-skewness and the ϵ addition addresses zero-inflation. This transformation improved RMSE by 2.1% beyond CARD alone but fell short of the ZI-HGT + CARD's 6.6% improvement. Clearly, a deterministic transformation cannot address ties in the OSCC ST data, nor does it allow UQ. HGT + CARD improved RMSE by 1.4%, emphasizing the need to address zero-inflation. Neither ALRA imputation nor MIST denoising + CARD significantly impacted performance. Median RMSE reductions were -0.1% and 0.2% respectively, underscoring the need for specific ST pre-processing techniques to transform the data to approximate normality. Other cell-type deconvolution methods performed notably worse than CARD alone; the best (SPOTlight) had 15% higher RMSE.

Finally, we examined the impacts of increasing the number of ties at small nonzero values in the data and altering sequencing depth on performance. To generate simulated ST data with additional ties at nonzero values, we randomly selected one cell type in the underlying simulated scRNA-seq data and added 1, 2, or 3 to the gene expression of a randomly selected 50% of genes, inducing ties at nonzero values in simulated ST data. The ZI-HGT improved performance beyond CARD alone by 7.8% when 1 was added, and 8.1% when 2 or 3 were added, demonstrating robustness to additional ties. To explore changes in sequencing depth,

we simulated ST data with -50%, -30%, -10%, +10%, +30%, and +50% depth and applied the ZI-HGT + CARD (Web Figure 48). The median improvement of the ZI-HGT + CARD vs. CARD alone is greater than 5% for the data with -10%, +10%, and +30% sequencing depth, and is approximately 4% with +50% sequencing depth, likely due to decreased sparsity (86.1%). Performance decreased when sequencing depth decreased dramatically (-30%, -50%). Decreased sequencing depth leads to lower signal, which is overwhelmed by the ZI-HGT's added noise. Thus, we do not recommend applying the ZI-HGT when ST data has few cells per location. Fortunately, ST techniques typically capture approximately 10 cells per location or achieve near single-cell resolution (Moses and Pachter, 2022).

5.2 Applying the ZI-HGT + CARD to CARD's Simulations

To demonstrate our method's viability in an unbiased manner, we applied the ZI-HGT + CARD to CARD's exact simulations without modifications. These simulations, constructed similarly to ours, are based on two freely available mouse olfactory bulb (MOB) datasets: a scRNA-seq reference (Zeisel et al., 2018) and an ST dataset (Ståhl et al., 2016). The simulated ST data is much less sparse than real ST datasets, with median sparsity of 68.6% versus 88.5% in the real OSCC ST datasets. Web Appendix B details the simulation procedure.

Five distinct simulation scenarios are considered: 1) The scRNA-seq data is applied as the reference basis matrix for cell-type deconvolution (standard approach). 2) One cell type is removed from the scRNA-seq reference. 3) One cell type (blood cells) is added to the scRNA-seq reference. 4) Two cell-types are merged in the scRNA-seq reference. 5) A mismatched scRNA-seq reference dataset from a different sequencing platform, mouse brain tissue, and mouse age is used (Zeisel et al., 2018). Following CARD, additional robustness comes from considering "noisy" data where the dominant cell type of a given percentage $p_n \in \{0\%, 20\%, 40\%, 60\%\}$ of locations is randomly chosen.

Figure 4 displays ZI-HGT + CARD's performance across 100 simulated ST datasets

for all five CARD simulation scenarios with $p_n = 0\%$ (Web Figures 5 - 7 show $p_n \in \{20\%, 40\%, 60\%\}$). The ZI-HGT + CARD is less impactful than in the realistic OSCC simulations (2.1% median RMSE reduction in scenario 1), which is expected given the unrealistically low sparsity of the CARD simulated dataset. For scenario 1 ($p_n = 0\%$), the median sparsity is just 68.6%, significantly lower than the 86%-91% in the OSCC data. As Figure 3 demonstrates, the ZI-HGT is more impactful in sparser datasets where zero-inflation is more pronounced. Overall, The ZI-HGT + CARD outperforms CARD alone for nearly every noisiness percentage p_n and scenario. The exception is scenario 5 with the highly mismatched scRNA-seq reference. As the ZI-HGT adds noise during transformation, the additional noise from the mismatched reference is amplified, slightly reducing accuracy. However, scenario 5 represents an uncommon, suboptimal edge case as using reference datasets from different tissues and platforms are generally avoided in practice. For mismatched references, using CARD alone is advised, and caution when interpreting results is encouraged due to increased risk of inaccuracy. We also compared the ZI-HGT + CARD to CARD alone applied to simulated data with $\pm 10\%$ sequencing depth (number of gene expression read counts per location) (Web Figures 4-7). As a sanity check, our analysis confirms that enhanced sequencing depth improves deconvolution accuracy.

6 Real Data Results

We applied the ZI-HGT + CARD to the OSCC data (Arora et al., 2023) introduced in Section 2. On all samples, we set the number of ZI-HGT iterations $C = 100$ and used the previously discussed HNSCC scRNA-seq data (GSE103322) as the deconvolution reference. We focus here on Sample 1 and present results for the Samples 2-12 in Web Figures 8-41.

OSCC Sample 1 contains spatially resolved gene expression for 15,844 transcripts across 1,131 locations. We analyze estimated proportions for all 14 cell types at each location in Figure 5 (ZI-HGT + CARD) and Web Figure 8 (CARD), displaying both proportion

estimates and binary plots with coloring applied only when estimated proportions exceed 0.05. We selected this threshold as 10x Visium ST data contains approximately 10 cells per location, and therefore proportions above 0.05 reasonably indicate that at least one cell of that type is present. Such presence or absence information is clinically valuable, as the presence of tumor-infiltrating immune cells (T cells, B cells, etc.) can predict OSCC patient survival (Hadler-Olsen and Wirsing, 2019). In practice, each cell type is not present at each location. Determining which cell types are absent is challenging as zero gene expression of marker genes can arise from multiple mechanisms: actual cell type absence, technical artifacts, or transcriptionally quiet cells due to environmental factors. This mechanistic uncertainty is reflected in the uncertainty of the low cell-type proportions estimates.

The ZI-HGT + CARD demonstrates marked improvement in identifying cancer-associated fibroblasts (ecm-myCAFs and detox-iCAFs) and normal fibroblasts (myofibroblasts and intermediate fibroblasts) across the tumor microenvironment (TME) compared to CARD alone. Normal fibroblasts can suppress tumor growth but can transform into cancer-associated fibroblasts during tumorigenesis (Maia et al., 2021). While both fibroblast types distribute throughout the TME, cancer-associated fibroblasts colocalize more closely with tumor cells (Maia et al., 2021)—a pattern clearly illustrated through the application of ZI-HGT + CARD (Figure 5a) but less evident with CARD alone (Web Figure 8a). Since cancer-associated fibroblasts are potential treatment targets (Maia et al., 2021), better determining their locations in the TME could improve targeted therapeutics.

Assessing the precision of deconvoluted cell-type proportions in real datasets is challenging, as the ground truth is unknown. A common metric calculates Pearson's correlation between mean estimated tissue-wide cell-type proportions and corresponding scRNA-seq proportions (Ma and Zhou, 2022; Li et al., 2023). For OSCC Sample 1, the correlation coefficient was 0.93 for the ZI-HGT + CARD vs. 0.85 for CARD alone (Figure 5). Across all 12 samples,

the ZI-HGT + CARD estimates showed stronger correlation with the scRNA-seq reference proportions for 10 samples, with notable differences in complex samples (1, 3, 4, and 5). While 42.1% of the cells in the scRNA-seq dataset are cancer cells (Puram et al., 2017), CARD alone estimates 90.0% cancer cells across all locations. The ZI-HGT + CARD also overestimates the overall proportion of cancer cells, but at a notably better 79.5%. Using the ZI-HGT clearly reduces CARD's substantial overestimation of dominant cell types.

Beyond CARD's capabilities, the ZI-HGT naturally quantifies uncertainty in cell-type proportion estimates through pointwise Bayesian credible intervals (Figure 6). For each cell type and location, we approximate the variance using the iterated variance formula and a second order Taylor series approximation, then construct pointwise credible intervals as the estimated proportion ± 1.96 standard deviations.

7 Conclusions

By integrating hierarchical generalized transformation models (Bradley et al., 2020; Bradley, 2022; Bradley et al., 2023) with Bayesian zero-inflation methods (Greene, 1994) and CARD cell-type deconvolution (Ma and Zhou, 2022), we address CARD's normality assumption and incorporate uncertainty in analyzing OSCC spatial transcriptomics data (Arora et al., 2023). The ZI-HGT is broadly applicable and can serve as an auxiliary technique for future ST analysis methods that assume normality.

The combination of the ZI-HGT and CARD yielded compelling insights into the spatial distribution of cell types within the OSCC TME. The ZI-HGT + CARD exhibited superior accuracy compared to CARD alone or with a naive transformation, enabling identification of diverse fibroblast populations—critical members of the TME affecting tumor growth and immunosuppression. A notable strength is the ZI-HGT + CARD's natural uncertainty quantification through the construction of pointwise Bayesian credible intervals without unrealistic distributional assumptions. UQ is rarely incorporated into ST analysis, and

we anticipate numerous applications the ZI-HGT's UQ framework. Researchers can use these credible intervals to assess the reliability of cell-type proportion estimates and apply pointwise Bayesian hypothesis tests to determine cell type presence at spatial locations.

Numerous methodological research directions remain open. New ST data capture techniques emerge annually, and adaptations of the ZI-HGT + CARD could extend to 3-dimensional single-cell resolution ST data. Approaches downstream of cell-type deconvolution, such as the cell-type-specific differential gene expression method C-SIDE (Cable et al., 2022), could be integrated into deep model structures with the ZI-HGT and CARD. Furthermore, techniques beyond cell-type deconvolution may benefit from the ZI-HGT, such as super-resolution gene expression prediction with iSTAR (Zhang et al., 2024). These promising ideas require substantial development and are reserved for future work. From a biological perspective, spatial patterns of estimated cell-type proportions warrant further exploration through collaboration with domain experts to fully realize these findings' potential.

We conclude by discussing potential limitations of the ZI-HGT + CARD. First, the ZI-HGT's transformations add a small amount of unavoidable noise to the ST data that could obstruct biological signal (Web Figure 45). This can be exacerbated when the sequencing depth is abnormally low (Web Figure 48). However, the ZI-HGT simultaneously reduces noise from CARD mis-specifying zero-inflated count-valued data as normal, yielding net improvement in deconvolution accuracy. The transformations may bias the data, as the high zero-inflation in the raw data results in insufficient "ground truth" to compare to. Thus, we used downstream deconvolution accuracy as a proxy to measure the effects of potential bias. Additionally, the limited integration of the ZI-HGT and CARD may be considered a weakness. CARD itself could potentially be altered to better align with the ZI-HGT and the data. CARD's conditional autoregressive (CAR) spatial structure may be too simple, particularly for non-stationary data where nearest neighbor adjacencies are unsuitable. A

more flexible spatial basis function expansion approach may prove superior, though this would require further exploration of the spatial structure of the ST data. Conversely, the independence of the ZI-HGT and CARD can be considered a strength. Alternative methods, such as the aforementioned iSTAR (Zhang et al., 2024), could replace CARD in the ZI-HGT + CARD model structure for tasks beyond cell-type deconvolution. We posit that the ZI-HGT provides a versatile foundation for future model development, enabling researchers to circumvent issues of zero-inflation and ties in ST data and simply assume normality.

Acknowledgments

We thank the editor, associate editor, and two anonymous referees for their constructive suggestions and valuable feedback on our manuscript.

Supplementary Material

Web Appendices, Tables, and Figures referenced in Sections 1-7 are available with this paper at the Biometrics website on Oxford Academic. All code used to generate the figures and results shown in this manuscript is available in the Supplementary Material.

Funding

H.J.M. is supported by the National Cancer Institute, National Institutes of Health grant 5T32CA134286-14. J.R.B. is partially supported by the U.S. National Science Foundation (NSF) grant DMS-2310756. C.W. is partially supported by the National Cancer Institute, National Institutes of Health grants U01CA293883 and R01CA263494.

Conflict of Interest

None declared.

Data Availability

The data used in this paper to illustrate our findings are made available by Arora et al. (2023) at `https://doi.org/10.6084/m9.figshare.20304456.v1` and Puram et al. (2017) at `https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE103322`.

References

Arora, R., Cao, C., Kumar, M., Sinha, S., Chanda, A., McNeil, R. et al. (2023). Spatial transcriptomics reveals distinct and conserved tumor core and edge architectures that predict survival and targeted therapy response. *Nature Communications* **14,** 1–19.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society Series B* **36,** 192–236.

Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society Series B* **26,** 211–252.

Bradley, J. (2022). Joint bayesian analysis of multiple response-types using the hierarchical generalized transformation model. *Bayesian Analysis* **17,** 127–164.

Bradley, J., Wikle, C. K., and Holan, S. H. (2020). Hierarchical models for spatial data with errors that are correlated with the latent process. *Statistica Sinica* **30,** 81–109.

Bradley, J., Zhou, S., and Liu, X. (2023). Deep hierarchical generalized transformation models for spatio-temporal data with discrepancy errors. *Spatial Statistics* .

Cable, D. M., Murray, E., Zou, L. S., Goeva, A., Macosko, E. Z., Chen, F. et al. (2022). Robust decomposition of cell type mixtures in spatial transcriptomics. *Nature Biotechnology* **40,** 517–526.

Danaher, P., Kim, Y., Nelson, B., Griswold, M., Yang, Z., Piazza, E. et al. (2022). Advances in mixed cell deconvolution enable quantification of cell types in spatial transcriptomic data. *Nature Communications* **13,** 1–13.

Elosua-Bayes, M., Nieto, P., Mereu, E., Gut, I., and Heyn, H. (2021). Spotlight: seeded nmf regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic Acids Research* **49,** e50.

Greene, J. (1994). Accounting for excess zeros and sample selection in poisson and negative binomial regression models. *NYU Working Paper No. EC-94-10* .

Hadler-Olsen, E. and Wirsing, A. M. (2019). Tissue-infiltrating immune cells as prognostic markers in oral squamous cell carcinoma: a systematic review and meta-analysis. *British Journal of Cancer* **120,** 714–727.

Lambert, D. (1992). Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics* **34,** 1–14.

Li, H., Zhou, J., Li, Z., Chen, S., Liao, X., Zhang, B. et al. (2023). A comprehensive benchmarking with practical guidelines for cellular deconvolution of spatial transcriptomics. *Nature Communications* **14,** 1–10.

Li, Q., Zhang, M., Xie, Y., and Xiao, G. (2021). Bayesian modeling of spatial molecular profiling data via gaussian process. *Bioinformatics* **37,** 4129–4136.

Linderman, G. C., Zhao, J., Roulis, M., Bielecki, P., Flavell, R. A., Nadler, B. et al. (2022). Zero-preserving imputation of single-cell rna-seq data. *Nature Communications* **13,** 1–11.

Ma, Y. and Zhou, X. (2022). Spatially informed cell-type deconvolution for spatial transcriptomics. *Nature Biotechnology* **40,** 1349–1359.

Maia, A., Wiemann, S., Reshkin, S. J., and Cardone, R. A. (2021). Cancer-associated fibroblasts: Implications for cancer therapy. *Cancers* **13,** 3526.

Markopoulos, A. K. (2012). Current aspects on oral squamous cell carcinoma. *The Open Dentistry Journal* **6,** 126–130.

Miller, B. F., Huang, F., Atta, L., Sahoo, A., and Fan, J. (2022). Reference-free cell-type deconvolution of multi-cellular pixel-resolution spatially resolved transcriptomics data. *Nature Communications* **13,** 1–13.

Moncada, R., Barkley, D., Wagner, F., Chiodin, M., Devlin, J. C., Baron, M. et al. (2020). Integrating microarray-based spatial transcriptomics and single-cell rna-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nature Biotechnology* **38,** 333–342.

Moses, L. and Pachter, L. (2022). Museum of spatial transcriptomics. *Nature Methods* **19,**

534–546.

Musgrove, D., Young, D. S., Hughes, J., and Eberly, L. E. (2019). *A Sparse Areal Mixed Model for Multivariate Outcomes, with an Application to Zero-Inflated Census Data*, pages 51–74. Springer International Publishing, Cham.

Puram, S. V., Tirosh, I., Parikh, A. S., Patel, A. P., Yizhak, K., Gillespie, S. et al. (2017). Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell* **171,** 1611–1624.

Ståhl, P. L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J. et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353,** 78–82.

Wang, L., Maletic-Savatic, M., and Liu, Z. (2022). Region-specific denoising identifies spatial co-expression patterns and intra-tissue heterogeneity in spatially resolved transcriptomics data. *Nature Communications* **13,** 1–12.

Watanabe, S. (2010). Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research* **11,** 3571–3594.

Yeo, I. N. and Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika* **87,** 954–959.

Zeisel, A., Hochgerner, H., Lannerberg, P., Johnsson, A., Memic, F., van der Zwan, J. et al. (2018). Molecular architecture of the mouse nervous system. *Cell* **174,** 999–1014.

Zhang, D., Schroeder, A., Yan, H., Yang, H., Hu, J., Lee, M. Y. Y. et al. (2024). Inferring super-resolution tissue architecture by integrating spatial transcriptomics with histology. *Nature Biotechnology* **42,** 1372–1377.

Zhao, P., Zhu, J., Ma, Y., and Zhou, X. (2022). Modeling zero inflation is not necessary for spatial transcriptomics. *Genome Biology* **23,** 1–19.

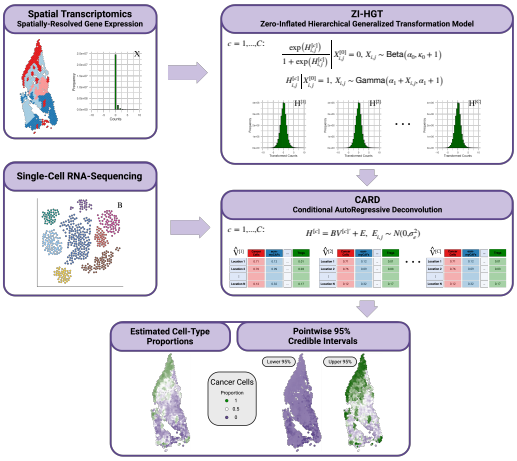
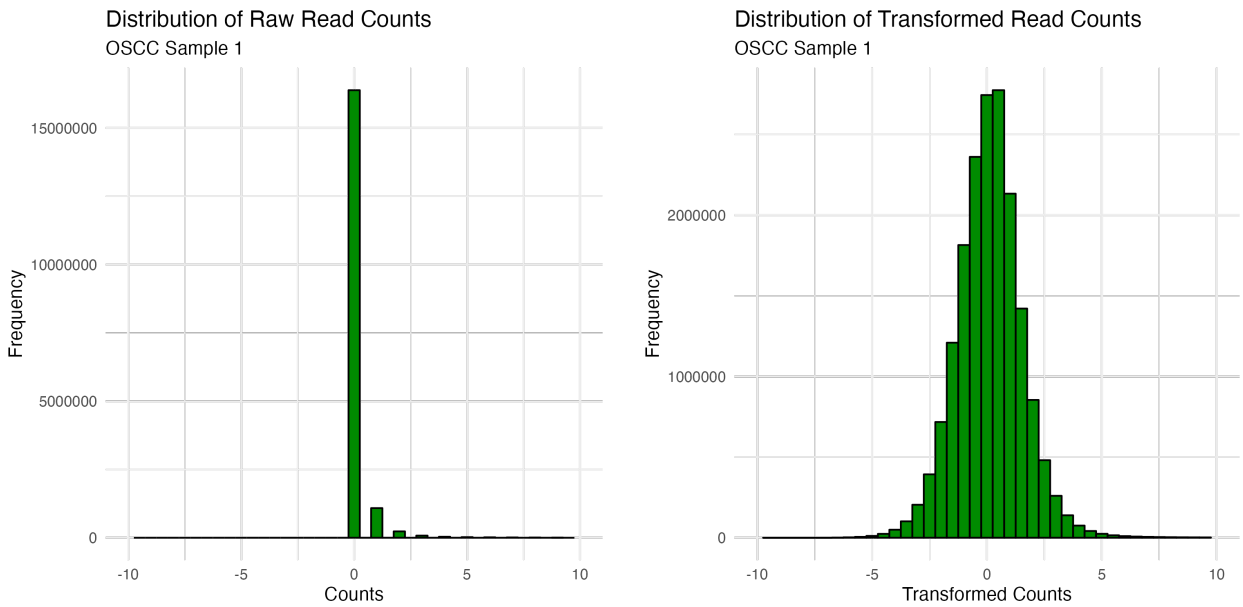


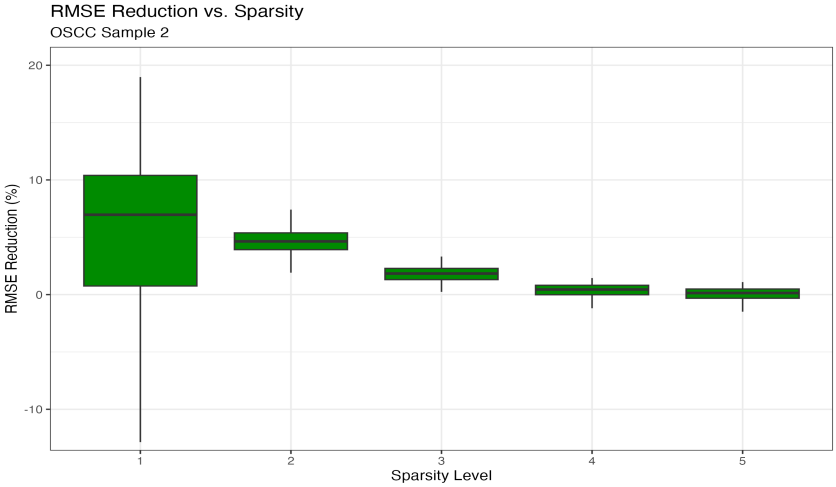
Figure 1: ZI-HGT + CARD Schematic Outline. The raw spatial transcriptomic data X is inputted into the ZI-HGT, generating C posterior replicates ($H^{[1]}$, ..., $H^{[C]}$), which are interpreted as transformations of the original data. These transformations better fit the assumptions of CARD, namely that the data are normally distributed. CARD is then applied to each transformed dataset and a corresponding single-cell RNA sequencing (scRNA-seq) dataset B, yielding cell-type proportion estimates ($\widehat{V}^{[1]}$, ..., $\widehat{V}^{[C]}$). These estimates are amalgamated into a single estimate $\widehat{V}$.



(a) Raw Read Counts

(b) Transformed Read Counts

Figure 2: Distribution of the gene expression read counts for OSCC Sample 1. a) Raw read counts. b) A single posterior replicate of the transformed read counts.



Sparsity Level	Φ Factor	Library Factor	Minimum	Median	Maximum
1	50	0.05	86.8%	89.8%	91.0%
2	10	0.1	73.2%	77.9%	80.1%
3	4	0.2	58.4%	63.7%	67.5%
4	4	0.4	48.8%	54.4%	58.3%
5	4	0.6	43.9%	49.7%	54.0%

Sparsity Level	Ties at 1	Ties at 2	Ties at 3	Ties at 4	Ties at 5	$\geqslant 5$
1	4.9%	2.0%	1.1%	0.6%	0.4%	1.3%
2	9.9%	4.5%	2.5%	1.6%	1.0%	2.6%
3	13.4%	7.2%	4.4%	3.0%	2.1%	6.4%
4	12.2%	7.3%	5.1%	3.7%	2.9%	14.4%
5	11.1%	6.9%	5.0%	3.9%	3.1%	20.2%

for all simulated ST datasets displayed in the table.

Figure 3: RMSE reduction vs. sparsity level of simulated ST data. RMSE reduction for the ZI-HGT + CARD across 100 simulated ST datasets and various sparsity levels. The sparsity levels in the boxplot have corresponding simulation hyperparameter factors, minimum, median, and maximum sparsity levels, and median proportions of data equal 1, 2, 3, 4, and 5 and $\geqslant 5$

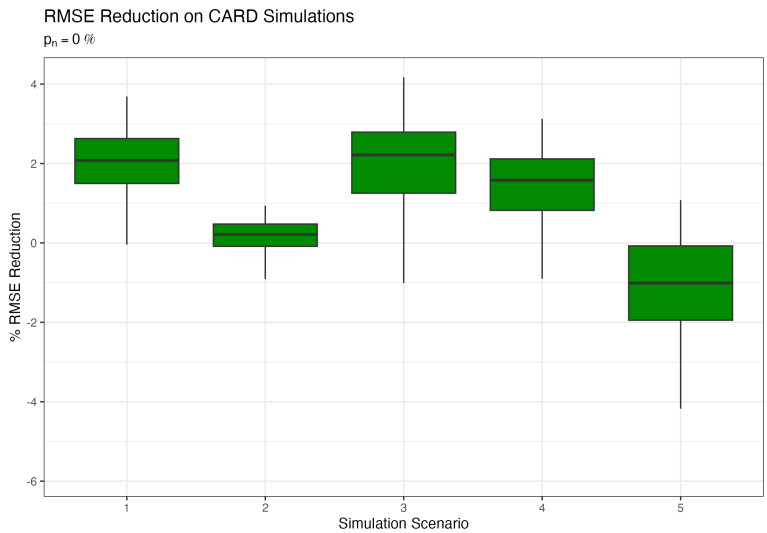
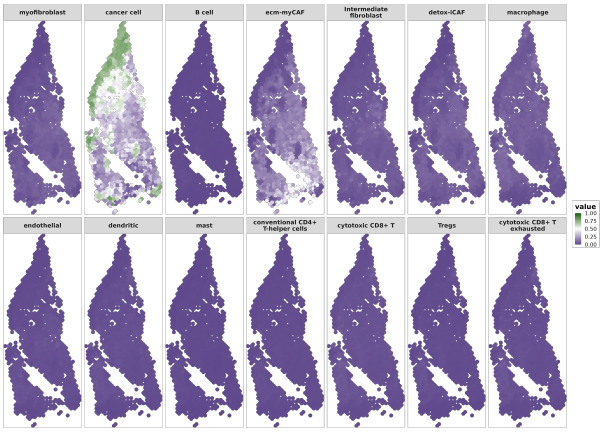
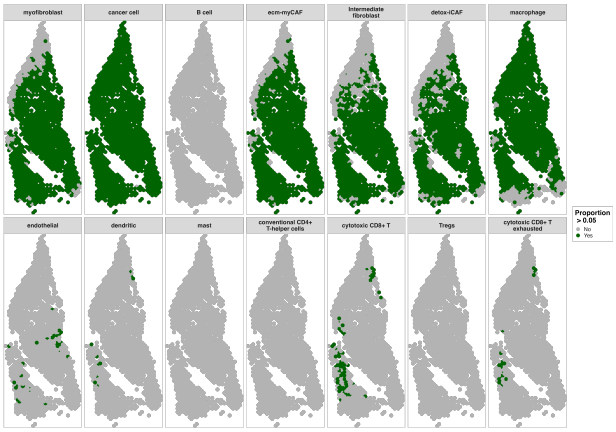


Figure 4: RMSE reduction on CARD simulations. RMSE reduction for the ZI-HGT + CARD across 100 simulated MOB ST datasets and various simulation scenarios following the CARD simulation setup with $p_n = 0\%$.



(a) Cell-Type Proportions Estimates

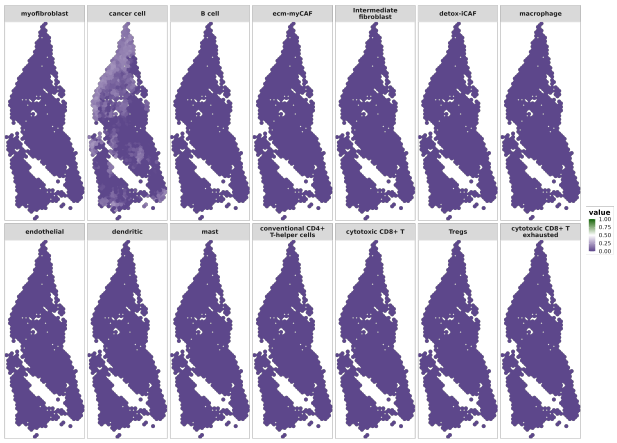


(b) Cell-Type Proportions > 0.05

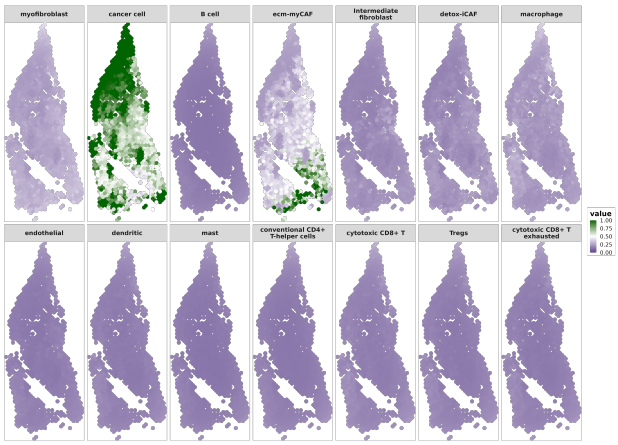
Cell Type	Reference	ZI-HGT + CARD	CARD
Cancer Cell	**0.421**	**0.421**	**0.494**
Myofibroblast	**0.124**	**0.058**	0.018
Conv. CD4+ Th Cells	**0.075**	0.019	0.001
ecm-myCAF	0.062	**0.158**	**0.256**
Cyt. CD8+ T Exhausted	0.047	0.026	0.009
Endothelial Cells	0.045	0.026	0.005
Detox-iCAF	0.043	0.061	0.032
Cyt. CD8+ T Cells	0.043	0.034	0.028
Tregs	0.038	0.020	0.003
Macrophages	0.027	0.067	**0.097**
B Cells	0.024	0.014	0.001
Mast Cells	0.021	0.014	0.001
Intermediate Fibroblasts	0.021	0.056	0.050
Denditic Cells	0.010	0.026	0.005

(c) Reference and Mean Cell-Type Proportions

Figure 5: OSCC Sample 1 ZI-HGT + CARD Cell-Type Proportions Estimates. a) Estimated cell-type proportions for all 14 cell types. b) Binary maps showing locations where cell-type proportions are > 0.05. c) Reference cell-type proportions from scRNA-seq and average estimated proportions across all locations using ZI-HGT + CARD and CARD. Values in bold indicate the three largest proportions for each method.



(a) Lower Bound



(b) Upper Bound

Figure 6: OSCC Sample 1 ZI-HGT + CARD Cell-Type Proportions Pointwise 95% Credible Intervals. a) Lower 95% bound. b) Upper 95% bound.