



Challenges and Opportunities in the Statistical Analysis of Multiplex Immunofluorescence Data

Christopher M. Wilson ¹, Oscar E. Ospina ¹, Mary K. Townsend ², Jonathan Nguyen ³, Carlos Moran Segura ³, Joellen M. Schildkraut ⁴, Shelley S. Tworoger ², Lauren C. Peres ², and Brooke L. Fridley ^{1,*}

- ¹ Department of Biostatistics and Bioinformatics, Moffitt Cancer Center, Tampa, FL 33612, USA; Christopher.Wilson@moffitt.org (C.M.W.); Oscar.Ospina@moffitt.org (O.E.O.)
- ² Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, FL 33612, USA;
- Mary.Townsend@moffitt.org (M.K.T.); Shelley.Tworoger@moffitt.org (S.S.T.); Lauren.Peres@moffitt.org (L.C.P.)
 ³ Department of Pathology, Moffitt Cancer Center, Tampa, FL 33612, USA; Jonathan.Nguyen@moffitt.org (J.N.); Carlos.MoranSegura@moffitt.org (C.M.S.)
- ⁴ Department of Epidemiology, Emory University, Atlanta, GA 30322, USA; Joellen.M.Schildkraut@emory.edu
- * Correspondence: Brooke.Fridley@moffitt.org; Tel.: +1-813-745-1461; Fax: +1-813-745-6107

Simple Summary: Immune modulation is considered a hallmark of cancer initiation and progression, and has offered promising opportunities for therapeutic manipulation. Multiplex immunofluorescence (mIF) technology has enabled the tumor immune microenvironment (TIME) to be studied at an increased scale, in terms of both the number of markers and the number of samples. Another benefit of mIF technology is the ability to measure not only the abundance but also the spatial location of multiple cells types within a tissue sample simultaneously, allowing for assessment of the co-localization of different types of immune markers. Thus, the use of mIF technologies have enable researchers to characterize patient, clinical, and tumor characteristics in the hope of identifying patients whom might benefit from immunotherapy treatments. In this review we outline some of the challenges and opportunities in the statistical analyses of mIF data to study the TIME.

Abstract: Immune modulation is considered a hallmark of cancer initiation and progression. The recent development of immunotherapies has ushered in a new era of cancer treatment. These therapeutics have led to revolutionary breakthroughs; however, the efficacy of immunotherapy has been modest and is often restricted to a subset of patients. Hence, identification of which cancer patients will benefit from immunotherapy is essential. Multiplex immunofluorescence (mIF) microscopy allows for the assessment and visualization of the tumor immune microenvironment (TIME). The data output following image and machine learning analyses for cell segmenting and phenotyping consists of the following information for each tumor sample: the number of positive cells for each marker and phenotype(s) of interest, number of total cells, percent of positive cells for each marker, and spatial locations for all measured cells. There are many challenges in the analysis of mIF data, including many tissue samples with zero positive cells or "zero-inflated" data, repeated measurements from multiple TMA cores or tissue slides per subject, and spatial analyses to determine the level of clustering and co-localization between the cell types in the TIME. In this review paper, we will discuss the challenges in the statistical analysis of mIF data and opportunities for further research.

Keywords: digital pathology; cancer; tumor immune microenvironment; data science

1. Introduction

Characterization of a patient's tumor immune microenvironment (TIME) is a topic that has stirred a great deal of interest with the advent of cancer immunotherapies [1,2]. Immunotherapies are agents that activate or act as a substitute for host antitumor immunity



Citation: Wilson, C.M.; Ospina, O.E.; Townsend, M.K.; Nguyen, J.; Moran Segura, C.; Schildkraut, J.M.; Tworoger, S.S.; Peres, L.C.; Fridley, B.L. Challenges and Opportunities in the Statistical Analysis of Multiplex Immunofluorescence Data. *Cancers* **2021**, *13*, 3031. https://doi.org/ 10.3390/cancers13123031

Academic Editor: Ewan Millar

Received: 14 May 2021 Accepted: 14 June 2021 Published: 17 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). and have been revolutionary in the treatment of many cancers [3,4], including melanoma, lung cancer, and renal cell carcinoma [5]. However, the efficacy in other solid tumors has been modest and is often restricted to a subset of patients [2,6,7]. Clinical trials investigating immunotherapy often lack indicators of immune activity, making it difficult to discern the characteristics of the tumor immune microenvironment (TIME), which may have impeded response to therapy. Thus, immune profiling has become an important tool to aid in our understanding of immune checkpoints and to identify predictive markers of therapeutic response.

For example, tumors with a dense infiltrate of lymphocytes, also known as tumor infiltrating lymphocytes (TILs), are consistently associated with more favorable outcomes among cancer patients [1,8,9]. In colorectal cancer, a novel immunoscore was developed and validated based on the density of TILs (CD3+, CD8+) within the tumor and invasive margin [10–12]. This immunoscore was strongly associated with cancer outcomes, showing superior promise as a prognostic marker compared to stage and microsatellite instability status and has provided a powerful prognostic tool to improve outcomes of cancer patients [13–16].

The cellular composition of the TIME can be studied using many technologies and approaches, such as immunohistochemistry (IHC), multiplex IHC (mIHC) [17], gene expression deconvolution methods (CIBERSORT [18], xCELL [19]), single cell RNA sequencing (scRNAseq) [20–23], flow cytometry [24], and approaches for mass cytometry imaging (imaging mass cytometry [25–27] and multiplex ion beam imaging (MIBI) [28,29]). Conventional IHC is a widely used technique in the field of diagnostic pathology [30,31]. This technique takes advantage of the epitope–antibody interaction to show in situ protein expression or biomarkers on a formalin-fixed and paraffin-embedded (FFPE) tissue sample [32,33]. More recently, immune-profiling by IHC is becoming an important tool to predict immunotherapy response in various types of cancer [34,35]. However, the inability to identify more than 2–3 markers per slide limits the use of IHC to capture the complexity of immune phenotypes that exist in the tumor microenvironment [33,36].

The development of multiplex immunohistochemistry (mIHC) and multiplex immunofluorescence (mIF) has allowed for the assessment of multiple markers in a single experiment [37]. mIF/mIHC can be applied to both regions of interest (ROIs) of a whole tissue slide or to tissue microarrays (TMAs), thus simultaneously allowing for the study of the TIME in large number of samples. Another benefit of mIF platforms that they can detect both the abundance and spatial location within the tissue sample of multiple cell types. The use of mIF has been recently applied to study the spatial proximity between T and PD-L1 expressing cells in oropharyngeal squamous cell carcinoma [38], the spatial heterogeneity of macrophages in gastric cancer [39], and the spatial composition of myeloid cells in pancreatic cancer [40].

There are many challenges in the statistical analysis of data from mIF (or mIHC) studies. First, many studies involve multiple cores or ROIs from the same tumor tissue sample (i.e., repeated measurements). Second, when studying tumors that tend to have little immune infiltration or "cold", often tumors have no positive cells for a marker of interest (i.e., zero-inflated data). To deal with this challenge, researchers often dichotomize the abundance measures into categories (i.e., no/low/high abundance). However, the challenge arises when defining the threshold to use in making these categories. Third, particularly in the case of TMA studies, there are often regions in which no cells were able to be measured (i.e., "holes" in the image). This uneven assessment of immune cells is often overlooked in spatial analyses, which often assumes that measurements are possible at all locations in the region of interest. Finally, batch effects between TMAs and phenotype misclassification (i.e., falsely calling a cell as positive for a marker) are common quality control issues related to mIF data. This review provides background on the analytic challenges related to mIF data, possible solutions and potential areas for future research. The concepts presented in this manuscript are illustrated using two large observational studies of ovarian cancer for which mIF data was recently generated (Nurses Health

Studies I and II (NHSI [41,42], NHSII [43]), the African American Cancer Epidemiology Study (AACES) [44]).

2. Data Preprocessing and Quality Control of mIF Data

2.1. mIF Assay and Data Generation

Multiple platforms exist for mIF techniques, including standard IF scopes and multispectral technologies (Vectra 3.0TM/PolarisTM). The most important step is the selection of the primary antibodies to target the biomarkers of interest, with monoclonal antibodies often being used due to their high sensitivity and specificity [37]. These antibodies are then labeled with fluorophores which emit wavelengths that can be measured via microscope with a corresponding image saved for image processing and analysis. mIF technologies allow for the use of multiple antibodies to achieve the simultaneous detection of several marker on single tissue sample (recently nine or more markers). This technology has been used in research and clinical settings showing the utility of this approach for studying the TIME [45-47]. An overview of the data generation process with mIF technology is presented in Figure 1, focusing on cyclic-immunofluorescence and tyramide-based mIF. Cyclic-immunofluorescence requires sequential cycles where individual epitope or markers are labeled with antibodies, which is then followed by a signal amplification. Individual antibody complexes are then stripped after each round of antigen detection leaving the fluorophore covalently attached to a tyrosine residue of the target epitope and ready for the next round of immunofluorescence [48]. This is a labor-intensive procedure and may take several days to complete. However, fully automated staining protocols for mIF have been developed, saving time and improving staining variability [45–49].

Specialized cameras and software are needed to properly acquire multiple markers in a single image. Multispectral imaging (MSI) is the main technology used to accurately capture mIF images, whereby the intensity wavelength spectrum of every pixel is captured in the image [50]. This procedure generates a third dimension of information for every pixel in the image and potentially increases the number of wavelengths that can be captured from 4 bands to 10–30 bands (multispectral). The information from each multispectral image pixel is extracted to correctly separate all the captured wavelengths per pixel and acquire the desired image [48,51–53] using a spectral reference or spectral library. Once the spectral library is built, images are spectrally unmixed and image files with channel marker metadata are processed with MSI analysis [54–58]. One commonly used technology for mIF data generation is the Vectra 3.0TM/PolarisTM system whereby images are processed within InForm [56,59–61] followed by analysis with the HALO Image Analysis Platform (Indica Labs, Albuquerque, NM, USA). In HALO, a supervised classifier using a random forest algorithm is trained to classify tissue as tumor, stroma, and glass (no tissue) regions [55,62,63]. Cell segmentation and marker quantitation is performed by compartmental examination of fluorescent intensity thresholds [60,61]. Analysis outputs are generated for each specific cell and summarized data per images including positive counts, co-localized phenotypes, marker intensity per compartment, percent of cells positive for a marker, and tissue area for density calculation and cell coordinates for spatial analysis [56,57,59,60].

2.2. Quality Control of Generated Data

2.2.1. Conflicting Information between Markers (CD8 and FOXP3)

In the process of cell phenotyping (i.e., calling a cell as positive for a marker), there are often situations in which cells are misclassified or mis-phenotyped. For example, in designing mIF assays, CD3 is often used as a general T-cell marker with additional markers added to distinguish between types of T-cells, with CD8 often used as a marker for cytotoxic T-cells and FOXP3 used as a marker for regulatory T-cells [64]. In the cell phenotyping, a machine learning algorithm is applied to the intensity data (i.e., machine learning algorithm within the HALO software) resulting in cells being called positive for a marker [56,65]. This approach may lead to equivocal cell type assignments, as illustrated in Figure 2A. In this case where cells were called positive for CD3, CD8 and FOXP3, one option would be to

classify the cells as only a T-cell (CD3+) and remove the conflicting assignment of cytotoxic T cell (CD8+) and regulatory T cell (FOXP3+). Another option would be to apply a different cell phenotyping algorithm to improve cell phenotype assignments, such as a supervised approach in which the highest intensity from a group of known false positives is used as the threshold for a given marker [66]. Finally, newer approaches have been developed to assist cell annotations, such as the CITE-Seq atlases [67].



Figure 1. (**A**) Data are generated from biopsied tissue that is FFPE preserved, slices are then placed on a tissue microarray (TMA). (**B**) The slide is stained with antigen which are the sites that primary and secondary antibodies bind. (**C**) A range of different wavelengths of light is radiated at each location of the specimen and the wavelength emission goes through a spectral unmixing step (**D**), which deconvolves the observed intensity into cyan fluorescent protein (CFP), yellow fluorescent protein (YFP), background, and Raman components. In order to phenotype each cell (**E**–**G**), the tissue is segmented into tumor and stroma component using staining (**E**), intensities and information regarding the shape of the cell is used to derive the final phenotype via machine learning (random forest is a popular technique), followed by cell phenotyping (**G**).

2.2.2. Batch Effects

Minimizing sources of assay variability (e.g., technical variation) is important for scientific validity, reproducibility and to maximize statistical power [68]. Sources of variability for mIF assays may include batch-to-batch variations across TMAs or a large set of whole slide ROIs and inconsistent quality of staining related to tissue characteristics, such as age of the tissue samples (e.g., if collected over years or decades). Therefore, after summary data have been generated for a set of samples (e.g., percent or density of positive cells), it is recommended that a quality control step be completed to compare the distribution of positive cells across subsets of samples defined by TMA, timing of staining or other technical "batch" factors, and/or tissue characteristics (e.g., histology, grade). Visual plots of the percent positivity, such as box plots, violin plots, and scatter plots, in each subset are useful tools for these qualitative quality control checks (Figure 2B). A binary positivity measure (e.g., percent of cases with >5% positive cells) could also be calculated. If certain tumor characteristics are associated with the prevalence of the cell type of interest and vary across sample subsets, these quality control checks should be conducted among similar samples (e.g., among cases with the same tumor histology).



Figure 2. (A) Square-root transformed CD8 (Opal 520) and FOXP3 (Opal 570) fluorescence intensities of a tumor microarray core from an epithelial ovarian cancer tumor. Cell classifiers used in immunofluorescence studies can yield equivocal CD8+FOXP3+ assignments. Note that the CD8 threshold creates a clear separation of CD8+ cells, however the FOXP3 intensity threshold allows for a mixture of unassigned and FOXP3+ cells. (**B**) Square-root transformed percent of CD8+ cells detected in 1312 epithelial ovarian cancer tumor slices from 445 participants. The tumor slices come from 6 different TMAs, with initial collection of tissues starting at different times since the 1970s. The three horizontal lines represent the 1st, 2nd, and 3rd quartiles, and the width of the violin plots represent the number of slices showing a given percentage. As showed by narrower violin bases, the TMAs generated starting in the 1990s show less zeroes in CD8+ cell counts compared to the other TMAs generated in previous years. (**C**) mIF images from the same core from an ovarian cancer TMA. The two slices were stained with pan-cytokeratin (PCK) but were applied two different mIF panels to detect B and T cells (top). The cells detected after image processing are shown. Differences are observed between the two slices, including the presence of "holes", making difficult to perform comparative spatial analysis of the two slices from the same TMA core. The white arrows correspond to a region that is similar across the different sections of the same core, while the green arrows correspond to regions that are dramatically different. Illustration that plots generated from mIF data capture these features and maintain the cell locations (bottom).

If the quality control checks reveal unexpected variability, the issues may be addressed by staining a new slide (or slides), returning to the image analysis stage, or conducting additional analyses during the statistical analysis. For example, examining distributions of markers across TMAs might reveal certain TMAs with much higher percentages of positive cases than others; this might be due to higher levels of background staining in those TMAs, which can be addressed by adjusting the image analysis. For a factor that cannot be altered, such as sample age, the analysis could adjust for this batch effect in the statistical model or conduct the analyses separately among the different subgroups of the factor of interest to examine whether the findings are consistent across subgroups. It is recommended that all cores from the same tumor be included on the same TMA so that the batch effect is not confounded with the within patient variability. It is also recommended when constructing TMAs for mIF studies that the use of randomization be done to ensure TMAs are balanced in terms of any relevant clinical (i.e., stage or tumor histology) and/or technical factor (i.e., age of tissue sample).

One of the most appealing features of immunofluorescence is the ability to assess numerous markers simultaneously, which allows for discrimination of a variety of cell types. Nonetheless, different panels of markers are applied to sequential tissue sections where the location of cells are not consistent. This use of sequential tissue sections makes different mIF panels not comparable in terms of the spatial organization of cells thus limiting the ability to study the colocalization of markers on different panels. Careful sequential cutting of the tissue to avoid rips or folding, while preserving orientation, requires a highly skilled and experienced technician [69], and it partially overcomes the challenge as cellular heterogeneity among sequential sections is expected (Figure 2C). It has been shown that different panels can be applied to the same tissue section, however, appropriate technical controls need to be carefully implemented to assess staining errors [53], and masking effects from previous markers can occur [70]. Consequently, spatial analysis needs to be conducted and interpreted separately by marker panel.

3. Analysis of Summary Data

3.1. Analysis of the Number, Percentage or Density of Cells Positive for Immune Marker

Frequently, the analysis of interest involves the extraction of summary statistics (i.e., does the abundance of PD-L1 positive cells differ between responders and non-responders to an immunotherapy). These summary statistics are calculated for each of the cell phenotypes and offer a general picture of the abundance across the tissue sample. Common summary statistics include the number, percentage, and the density of positive cells expressing a given marker. These metrics can be calculated manually, although higher accuracy and efficiency can be achieved by using image analysis software (e.g., HALO) [71]. To account for TMA or ROI size, the percentage of positive cells out of a total number of measured cells and/or counts per unit of area (density) are widely used. These summary measures can be computed separately for the tumor and stromal compartments [72]. When measuring immune cells in "cold tumors", the total number of immune cells is expected to be a very small proportion of the total number of cells, possibly leading to apparently small between person variation (Figure 3A). For especially rare immune cell subsets, an alternative approach would be to look at the percent of a type of immune cell out of only the immune cells, not out of the tumor and stroma cells (e.g., number of CD8+CD3+ cells out of total number of CD3+ cells). These summary measures are then used in statistical models to determine associations between immune markers and various outcomes. However, care in picking the statistical modeling approach is needed as often these summary measurements do not follow a normal distribution.

Many statistical analyses approaches are based in Gaussian or normal distribution theory for parameter estimation (confidence intervals) and hypothesis testing (*p*-value). For the analysis of the number of positive cells (X_i) out of the total number of cells measured (N_i) for sample *i*, the distribution for the analysis is a binomial distribution (i.e., $X_i \sim$ $BIN(N_i, p_i)$). If the probability of being a positive cell (p_i) in the sample is low (i.e., "rare event" or "cold" tumor) and assuming the number of cells (N_i) is large, the binomial distribution can be approximated with a Poisson distribution (i.e., $X_i \sim POI(\lambda_i = N_i p_i)$). In these distributional settings, a generalized linear model can be used to assess the association of a predictor variable on the number of positive cells for a given marker (i.e., does the level of cell positivity differ between treatment groups) [73,74].

Another approach often used is to model the percentage of positive cells (X_i/N_i) or the density using normal theory methods (i.e., two-sample t-tests, linear regression, ANOVA) or with the corresponding non-parametric method (i.e., Wilcoxon rank-sum test or Mann–Whitney U test). Prior to using the normal theory methods, a plot of the percentages or densities by the predictor variable (i.e., treatment group) as well as a plot of the residuals from the statistical model should be conducted to determine if the normal distributional assumption is reasonable [75,76]. As the percentage is bound between 0% and 100% and for immune cells, many samples have few cells positive for the marker of interest, a left skewed distribution or "pile-up" of observations by 0% (Figure 3B) is typical. In this case, a suitable transformation to the percent or density, such as a natural log (ln), logit (i.e., $\log(p/1 - p)$), or square-root transformation, should be applied. Note that in the setting in which the dataset contains 0% (or 100%), only a square-root transformation is possible.



Figure 3. (**A**) Illustration of the possible differences in immune activity within TMA cores. (**B**) Histograms with empirical and theoretical probability density function (top) and empirical and theoretical cumulative probability distribution (bottom) to guide in the selection of modelling assumptions for markers becoming increasingly rare (from left to right). The Poisson and binomial distribution do not account for overdispersion or zero-inflation, negative binomial and beta-binomial only account for over dispersion, and zero-inflated Poisson and binomial distribution only account for zero-inflation. The negative binomial and beta-binomial distributions are suitable for cell types where less than 50% of the cores have 0 for that cell type (CD3+, CD138+), while zero inflated models are best for excess 0s (CD19+).

Alternatively, the continuous measure of abundance (e.g., percentage positive, density) can be categorized into groups. The challenge with this approach is the selection of the threshold used to make the categories. Often, researchers use the median (50th percentile) or the quartiles (25%, 50%, and 75%) of the values in a dataset to construct the various groups/categories. In the context of biomarkers, often a biologically relevant threshold is selected. For example, in breast cancer, the estrogen receptor (ER) must be expressed in more than 1% of cells to be called ER+ [77]. One challenge is determining the clinically and biologically relevant thresholds for "positivity". For example, it is difficult to compare

results between clinical trials for PD-1/PD-L1 as studies have used different thresholds for determining who received benefit from immunotherapies, with some trials using a threshold of 1% positivity in tumor cells [78], while other studies have used a threshold of 10% [79]. This threshold for PD-L1 "positivity" can also differ from 1% to 50% in different cancer types due different biological and immune mechanisms [80–82], with little research presented on thresholds for other immune markers, such as cytotoxic T-cells.

An alternative approach for setting the pre-defined threshold is to determine the "optimal cut-point", which is selected to maximize the test statistic of interest [83,84]. From a statistical standpoint, the optimal cut-point approach is "data-snooping" since the results from the statistical test inform the creation of the thresholds [85]. On the other hand, for discovery and hypothesis generation purposes, it is clinically and biologically useful to determine an optimal cut-point that can be used and validated in other studies. However, a challenge in determining optimal cut-points is that the number of samples in a group/category can get very small when categorizing across multiple variables. In a study of the TIME of ovarian cancer in African American women, the optimal cut-points as related to overall survival for dichotomizing CD3+, CD3+CD8+ and CD3+FOXP3+ ranged from 1–6%, restricting groups to have at least 10% of the sample size [86].

3.2. Analysis Using Zero-Inflated and Over-Dispersed Distributions

An additional challenge in the analysis of summary data from mIF studies are scenarios in which many of the samples have no positive cells for a given biomarker of interest (i.e., zero-inflated distribution; Figure 3B). This is especially true for "immune cold" tumors where only a few cells express the immune markers of interest. In recent decades, extensive research has been conducted in zero-inflated statistical modeling, mainly for a Poisson distribution [87], referred to as a "ZIP" model. Zero-inflated models have been recently extended to the analysis of microbiome or metagenomics data, either in the context of Poisson [88], negative binomial [89,90], or beta-binomial [91] distributions. These models have also been extended with random effects to account for repeated measurements or dependency in observations [87,88,92].

In general, the zero-inflated models are statistical methods that allow for frequent zerovalued observations or "overdispersion". These models are considered a type of "mixture model" or "two-part model" where the model involves a mixture of a standard distribution (i.e., Poisson) and a point mass distribution at 0. Assuming the count of positive cells follows a Poisson distribution (as an approximation to the binomial distribution), another approach to account for overdispersion is with a negative binomial distribution. For example, in the context of RNA-sequencing data analysis, the read counts from high-throughput sequencing often are assumed to follow a negative binomial distribution instead of a Poisson distribution to account for the overdispersion in the data (i.e., the variance in gene expression abundances is much larger than the mean abundances, a departure for a Poisson distribution which assumes the mean equals the variance) [93–95]. Another option to account for the zeroinflation would be to model the data with a beta-binomial distribution. The beta-binomial distribution is the binomial distribution, $X \sim Bin(N, p)$, in which the probability p of being a cell positive for an immune marker out of N cells is not fixed but rather a random variable with $p \sim Beta(\alpha, \beta)$. The beta-binomial distribution has been extensively used in biological and medical research [96–101], frequently within a Bayesian framework [102]. As no one distribution will fit all markers in a study the best, it is recommended to assess model fit for the various distributions at the beginning of the analysis and select the appropriate model for each marker of interest, thus providing the highest power to detect true associations (Figure 3B). In our experience, the zero-inflated binomial models often fit best when there is a larger percentage of 0 positive cells counts while the beta-binomial works well when the level of zero counts is not as extreme.

3.3. Repeated Measurements

mIF studies often include multiple tissue samples from the same tumor (i.e., multiple cores in the case of TMAs or regions of interest (ROIs) in the case of whole slides). This aspect of the data requires the application of statistical analysis methods that account for dependency in measurements taken from the same subject. Often, this is completed via the use of mixed- or random-effect modeling [103] within the context of a linear model (i.e., linear regression using a normal distribution) or generalized linear model (i.e., logistic regression using a binomial distribution). The challenge in fitting these mixed models, particularly in the case of a generalized linear mixed model, is that they can be computationally intensive and require numerical approaches for parameter estimation and hypothesis testing, such as use of the expectation–maximization (EM) algorithm [104], Newton–Raphson method [105], numerical and Gaussian quadrature, or Markov chain Monte Carlo (MCMC) [106,107]. Hence, one of the key steps in fitting mixed models is checking for convergence of the estimation algorithm.

Alternatively, if one is not able to fit these more sophisticated mixed models, the mean of the density or percent positivity across the samples from the same subject can be used. However, this approach may under-estimate the variability in the measurements (except in cases where the repeated values for a subject are very consistent) and hence could impact the statistical inference (i.e., the type I error rate will be increased and the confidence intervals will be too small [108]). However, for very small studies, this might be the only approach possible. Lastly, if you assume the density or the percentage of positive cells follows a normal distribution and are interested in whether these values differ between two or more groups, a repeated measures ANOVA analysis could be completed; noting that unlike mixed models, additional covariates or predictors cannot be included in the model [109].

4. Clustering and Cooccurrence in Spatial Analysis of mIF

Research in ecological and spatial statistics over the last 40 years have developed many analytical methods that can be leveraged for studying the spatial architecture of the TIME. As such, many of these methods have been recently applied to the analysis of mIF data (Table 1). In general, these methods can be applied at the pixel/region-level or the cell-level and can be classified into three collections: count based methods at the pixel or region-level; point process methods at the cell-level; and distance-based methods at the cell-level. The pixel or region-level methods typically quantify a measure of interest for each pixel, followed by assessment of the variation in measurements in order to study the diversity, heterogeneity or autocorrelation across the entire image. On the other hand, cell-level methods often involve studying the nearest neighbor of each individual cell or the distance between cells to quantify the degree of clustering, cooccurrence, or segregation of cell populations.

Type of Analysis	Name	Empirical Formula	Theoretical Value under CSR	Comments
Pixel/Area Based	Morisita Horn Index [110,111]	$MH(p_1, p_2) = \frac{2p_1p_2}{p_1^2 + p_2^2} = \frac{2\sum_{k=1}^{p} p_1^k \times p_2^k}{\sum_{k=1}^{p} (p_1^k)^2 + \sum_{k=1}^{p} (p_2^k)^2}$		• Robust to settings involving small number of cells [112]
	Duncan Segregation Index [113]	$D = 2^{-1} \sum_{k=1}^{p} \left p_1^k / p_1 - p_2^k / p_2 \right $		Do not work well for rare cell populationsCheckerboard Problem [114]
Nearest Neighbor	Euclidean Distance	$d(c_i, c_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$	$\left(\lambda\pi r^2\right)^{-1}$	
	Nearest Neighbor	$\min_j d(c_i, c_j)$	$\left((n-1)\lambda\pi r^2\right)^{-1}$	

Table 1. Summary of the spatial measures outlined in Section 4 with the distinction for the spatial point processes being made to highlight the duality between distance to the nearest neighbor and locations of events.

Type of Analysis	Name	Empirical Formula	Theoretical Value under CSR	Comments
Spatial Point Processes	Ripley's K [115]	$\hat{K}(r) = (n(n-1))^{-1} \sum_{i=1}^{n} \sum_{i \neq j} 1(d(c_i, c_j) \le r) e_{ij}$	πr^2	• Summarizes larger scale than G
	Besag's L [117]	$\hat{L} = \sqrt{rac{\hat{k}(r)}{\pi}}$	r	Can be modified to handle non-homogeneous
	Marcon's M [118]	$\hat{M} = rac{\hat{k}(r)}{\pi r^2}$	1	 Cumulative and no information on what radius the clustering occurs Better overall performance than F, G, J [116]
	Pairwise Correlation Function [119,120]	$\hat{g}(r) = \left(2\pi\right)^{-1}\sum_{i=1}^{n}\sum_{i\neq j}rac{\kappa\left(r-d\left(c_{i},c_{j} ight) ight)}{d\left(c_{i},c_{j} ight)}e_{ij}$	$\frac{K'(r)}{2\pi r}$	 Not cumulative Interpreted as the probability two cells are a specified distance apart Best suited for cells with signaling processes
	Hypothesized Interaction Distribution [121]	$\hat{h}(i, j) = n^{-1} \sum\limits_{i=1}^n \sum\limits_{i eq j} 1 ig(c_i, c_j ig) \leq r ig)$	$(n-1)*\pi r^2$	No edge correctionMean increases with number of cells
	Empty Space Function [122]	$\hat{F}(r) = m^{-1} \sum_{i=1}^{m} 1\left(r \le \min_{j} d\left(l_{i}, c_{j}\right) \le r + \Delta r\right) e_{ij}$	$1 - \exp(-\lambda \pi r^2)$	 Summarizes much smaller scale than K, L, and M For a point process, both G and F have the same distribution Sensitive to processes with distance between point constraints [116]
	Nearest Neighbor Function [116]	$\hat{G}(r) = n^{-1} \sum_{i=1}^{m} 1\left(r \le \min_{j} d(c_i, c_j) \le r + \Delta r\right) e_{ij}$	$1 - \exp(-\lambda \pi r^2)$	
	Hazard Empty Space Function [123] or Hazard Nearest Neighbor Function	$h_{lpha} = rac{d}{dr} \left(-\log(1 - \hat{lpha}(r)) ight)$	$2\pi r\lambda$	Interpretation similar to time to event survival analysis
	J-function [124]	$\hat{f}(r) = rac{1-\hat{\mathcal{L}}(r)}{1-\hat{F}(r)}$	1	 Constant mean value Robust against edge corrections

Table 1. Cont.

P = number of pixels; p_i^i = proportion of the population of cell type *i* in the *j*th area or pixel; p_i = proportion of the population of cell type *i* across the entire image; n = total number of cells; e_{ij} = edge correction for the *i*th and *j*th cell; l_i = the *i*th randomly selected location; κ is a kernel function; h_{α} = hazard function of α = *F* or *G*. Blue text corresponds to spatial point processes that based on the location of cells. These methods are also referred to as second order methods. Red text corresponds to spatial point process methods that focus on distance to the nearest neighbor.

There are challenges in using these co-localization or clustering measures directly due to image curation issues (i.e., "holes" or areas in the image with no cells measured particularly in the setting of TMA studies) (Figure 4A) and/or studying rare cell types (large proportion of samples with 0 positive cells) (Figure 3A), which may lead to departures from the underlying assumptions for which the statistics were developed. In order to relate co-localization or clustering of immune cells to a phenotype (i.e., survival, treatment response), it is also necessary to ensure that data collected across multiple samples is comparable by (1) normalizing measurements across samples; (2) correcting the estimate to account for regions in the image in which no cells were able to be measured; (3) correcting for edge/border effects; and lastly (4) accounting for the correlation between abundance and spatial measures (i.e., not able to estimate spatial clustering when the abundance is close to 0%). These steps are critical to ensure the magnitude of features have the same meaning across the TMA cores or ROIs.

One way to account for many of these issues is with the use of permutation or Monte-Carlo methods, where the status (i.e., positive for immune marker) of independent blocks of pixels or cell is permuted, thus providing a broad set of scenarios to help evaluate the significance of the observed pattern in the image [125]. Using permutations, image specific distributions of the statistic of interest are estimated under the assumption of complete spatial randomness (CSR) (i.e., null distribution). The mean of this estimated distribution under CSR can then be used to normalize the estimate for each sample, thus producing a measure of the "degree of spatial clustering or colocalization" (i.e., observed spatial statistics—the mean of the empirical distribution of the statistic under the assumption of CSR) (Figure 4E). The association analyses would be based on this degree of clustering statistic and an adjustment for the overall abundance of the marker in determining the involvement of spatial co-localization/clustering on the phenotype of interest (i.e., response to immunotherapy, overall survival, tumor grade). The permutation of independent blocks of pixels or cell locations can provide a broad set of scenarios to help evaluate the significance of the observed pattern in the image, allowing for assessments separately in the tumor and stroma compartments of the tumor [125].



Estimate Observed K Permuted CSR Theoretical CSR

Figure 4. (A) Example images showing cores with little to significant damage. A point process generated from simulated data illustrating different approaches for spatial analysis: (B) Distanceor nearest neighbor-based methods, (C) neighborhood methods such as K(r), L(r), and M(r); and (D) distance to neighbor measures such as h(r) and g(r). (E) Example of original (left) and permuted point process (middle) with resulting histogram (right) of permutation-based estimates of *K* showing difference in theoretical and permuted-based estimates of CSR where the theoretical value under-estimated the value of *K* under CSR.

4.1. Pixel or Region-Based Methods

The first type of methods used in studying the spatial contexture of the TIME are pixelor region-based methods. These methods involve splitting a TMA core or ROI into separate non-overlapping regions where a summary measure of two cell types can be computed (e.g., number of positive cells; number of cell types observed in region). One such method is the Morisita-Horn (MH) index, which is a measure of spatial dispersion of individual populations of interest (i.e., immune cell populations) [110,111]. The MH index was recently used in a study of HER2+ breast cancer, in which a high degree of colocalization between immune and tumor cells was associated with a higher probability of 10-year survival [126,127]. However, a limitation of the MH index is that this measure does not provide a reliable estimate when one of the immune cell populations is rare, with methods proposed to reduce the under-sampling bias [128]. In contrast, the Duncan segregation index [113], developed in the context of gender based occupational segregation, can also be used as a region-based segregation or co-localization measure to determines if the proportion of immune cell populations in a region (i.e., CD3+CD8+ cells vs. CD3+FOXP3+ cells) differ from the expectation under no co-localization. However, neither MH index nor Duncan's segregation index account for the spatial autocorrelation between neighboring regions that is known to be the case with cellular processes, such as cell signaling. Cell signaling is one mechanism by which T cells are regulated, and it has been discovered that CD4 and T-cell antigen receptor (TCR) cells tend to form separate clusters in protein islands while CD4 and TCR cluster together upon T cell activation [129].

In contrast to the MH and Duncan segregation index developed for ecological and social science research, the Voronoi diagram [130] was developed in the field of mathematics for the study of quadratic forms and has been used extensively in geophysics and meteorology for spatial analysis. In terms of application to study of the TIME, a variation of Voronoi diagrams was used to measure "cell sociology" in lung adenocarcinoma [131]. Using the neighbors defined by the Voronoi diagram, cell-cell interaction measures were computed based on "adjacency" with permutations to obtain an image specific null distribution. This method was used to reveal that high T-cell clustering is associated with lack of recurrence. Additionally, in non-recurrent cases, a higher frequency of tumor cells with neighboring CD3+CD8- cells were observed than expected by chance [131].

4.2. Distance- and Nearest Neighbor-Based Methods

One of the most common methods to study the co-localization and cooccurrence of immune cells in the TIME is nearest neighbor distance (NND) [60,132,133]. This approach can be used to determine which immune cell type tends to cluster close to tumor cells by computing the distance (e.g., Euclidean) between the immune cell and closest tumor cell (Figure 4B). NND was used in a study of pancreatic ductal carcinoma where it was observed that myeloid cells (CD16+) were closer to tumor cells than T and B cells [134]. In melanoma, NND analysis showed that proximity of cancer cells to cytotoxic lymphocytes depends on the expression of Ki67 (tumor cell proliferation marker) and that the expression of HLA-DLR (macrophage activation marker) impacts the proximity of macrophages to cytotoxic lymphocytes [132]. However, careful attention to the implementation of this approach is needed when large "holes" or regions are present in TMAs where cells are not able to be measured. To overcome this challenge, permutation or Monte-Carlo methods can be used, as outlined in the next section.

4.3. Spatial Point Process Based Methods

The final type of spatial analysis methods for studying the TIME are those that are based on spatial point processes. In this setting, the locations of the immune cells in the TIME can be thought of as a spatial point process, a random pattern of points in a predefined area or "window". The simplest point process is a homogeneous or stationary Poisson point process, where the rate of an immune cell is constant over the entire region of interest [116,135,136]. However, in studying the TIME, we often wish to know if the

arrangement of cells positive for an immune marker fail to meet the assumption of CSR for homogenous spatial processes. Attraction (clustering), repulsion and competition (dispersion) are examples of interactions that would lead to violation of CSR.

Analysis of spatial Poisson point processes can consist of studying the number and location of points within a certain region, following a Poisson distribution. An alternate way to study spatial Poisson point processes is in terms of studying the spacing of events (distance to nearest neighbor) to understand how events are clustered, where the distance of the nearest neighbor follows an exponential distribution. Several quantities will be described in the following sections (Table 1), but it is important to remember that these concepts are complementary.

4.3.1. Analyzing Number of Neighbors

Clustering and cooccurrence can be studied directly by analyzing the location of cells themselves. This requires defining a neighborhood as circle, with a specified radius r, surrounding a cell and defining each cell within the neighborhood (Figure 4C). These methods can be applied to questions related to a single cell type or two cell types with the assumption that the underlying point process follows CSR. A popular quantity that studies the number of nearest neighbors and has been used in ecological statistics for the clustering of objects (i.e., trees) is Ripley's K(r) [115,116], where K(r) is expected to increase quadratically with respect to r. Besag introduced a modification to K(r), L(r), which theoretically increases linearly with r [117]. The degree of clustering is estimated as the difference between the observed estimate and the estimate under CSR, where a positive difference indicates clustering and a negative distance indicates dispersion. Marcon also proposed a modification to K(r), referred to as M(r), where the expected value is 1 for all values of r and can be interpreted as the percent of clustering/repulsion observed [118]. Similarly in concept to K(r) and increases quadratically with respect to r, the hypothesized Interaction Distribution (HID) has been proposed to measure the interaction or co-localization between immune cells [121]. However, the HID does not incorporate edge corrections or centers the observed quantity about the expected value. The application of HID to oropharyngeal squamous cell carcinomas found that high co-clustering of CD8+ and PDL1+ as well as CD8+, PD1+ and PDL1+ was associated with worse survival [38].

4.3.2. Analyzing Distance to Neighbor

The quantities in this section have a different interpretation than the neighborhoodbased measures. They are interpreted as probabilities of an event occurring within a radius r. The nearest neighbor distance distribution ("event-to-event" distribution) is denoted by G(r) [116]. A counterpart to G(r) is the empty space function, F(r) [122], or often referred to as the spherical contract distribution (SCD), which is estimated by selecting arbitrary locations as opposed to the search region being centered at each cell. Similar to K(r), it is computed as a variety of radius values, r, and can be compared to the estimate under CSR by direct comparison to the theoretical distribution or computing J(r). Bivariate versions of F(r) and G(r) have been derived to study the spacing between two cell types and has been used as a surrogate for immune cell infiltration of a tumor, where interactions between T-regulatory cells (Tregs, CD3+FOXP3+) and tumor cells, and cytotoxic (CD3+CD8+) and Tregs were shown to improve survival in patients with NSCLC lung cancer [137].

Finally, since the G(r) is the cumulative distribution function (CDF) of an exponential random variable, a hazard function can be estimated with the interpretation as the chance the nearest neighbor is within a small ring (Figure 4D). The pair correlation function or radial distribution function, g(r) and J(r), respectively, and the spherical contact distribution (SCD) or F(r) were used to study CD68+ macrophages in human head and neck tumors [119]. In this study, they found that the level of clustering of CD68+ cells (macrophages) is related to the level of leukocyte infiltration in the sampled ROI. The function g(r) describes how the density of cells differs as a function of distance r from a reference cell [120] and has been used extensively in the field of physics [138], while J(r) involves the ratio of 1 - G(r) to 1 - F(r) and its expectation is 1 under CSR [124].

4.3.3. Considerations

One challenge in using point process methods is the correction for edge effects. An edge effect arises as cells on the border lose neighboring cells that are located outside the sampled region. Corrections for edge effects have been developed for many methods, including Ripley's K(r). Two common adjustments are isotropic and translation corrections [135,139]. These corrections are necessary as the point process continues outside of the study region; hence analysis using only the cells measured in the image will underestimate the number of cells in the proximity for cells near the border. An excellent demonstration and explanation of edge corrections are in Gabriel et al. [140].

Another issue in using point process methods is that they often involve functions computed at a variety of radii values (r) producing a function or curve. One approach would be to select a value of r to use to estimate the spatial measures that would then be used in the association with the phenotype of interest. Alternatively, the spatial measures could be treated as all the values of r as a function or curve and complete functional data analysis (FDA). FDA is the analysis of information in curves [141]. One approach would be to compute the area between the empirically observed function/curve, where the curve would be computed under the assumption of CSR. This area, as opposed to the curve, can be used for the association analysis. This approach was recently used to characterize the immune landscape of malignant pleural mesothelioma tumors into two distinct patterns related to the clustering of tumor-associate immune cells [133].

We have applied Ripley's K(r) to the context of ovarian cancer, where patients with tumors having high abundance and a low degree of spatial clustering of CD3+CD8+ cells had significantly better survival compared to the patients with high spatial clustering [86]. A permutation approach was used to estimate the distribution of K(r) under CSR in order to correct for the potential bias in measurements due to "holes" in the TMA cores (Figure 4E). Additionally, edge effects were corrected and with the use of permutation or Monte-Carlo methods, clustering was assessed in the tumor and stroma components separately.

5. Discussion and Conclusions

In this review paper, an overview was provided of the statistical challenges and opportunities in the analysis of mIF data ranging from data quality control and assessment of batch effects to the statistical analysis with a zero-inflated distribution within a repeated measurements context. The mIF technology allows for an in-depth analysis of the tumor immune microenvironment (TIME) with assessments of both the abundance and spatial location of immune cells in the tumor microenvironment. The described analytical approaches are also applicable for the analysis of data from other technologies used to study the TIME (i.e., mIHC [17,37] and image mass cytometry [28]). Additionally, we addressed challenges of analyzing data from both TMAs and ROIs. In particular, TMAs tend to have smaller tissue areas (i.e., fewer cells measured) and areas that are not able to be assessed (i.e., "holes" in the image).

The presented analysis methods assumed that cell phenotyping of tissue specimens had been completed by a machine learning algorithm (i.e., converting the intensity measurement to a binary call of positive or negative for a cell type marker). Another approach would be to complete the statistical analysis on the raw intensity data, thus avoiding the issue of cell misclassification. However, care is needed as it is not clear if the intensity values have any intrinsic meaning. Additionally, batch effects in the intensity values are possible and additional analysis is needed to normalize and correct for batch effects in the samples [142–144].

For the analysis of the summary measures, it is recommended to assess model fit to determine the appropriate testing framework (i.e., use of zero-inflated or over-dispersed distribution) and if possible, to model using a count-based distribution (i.e., binomial

distribution). In the context of survival analysis where the focus is assessing the association of the abundance of an immune marker with a time to event endpoint (i.e. overall survival, time to progression on immunotherapy), care is needed in selecting the modeling approach. When there are many samples with zero positive cells and/or a very skewed distribution, the abundance measure (i.e., percentage, density) should be categorized based on biologically relevant cut-points to aid in the interpretation and comparison of results across studies.

A benefit of using mIF technologies is the generation of spatial information for each cell. In studying the spatial architecture of the TIME in many samples/subjects, the spatial measurements should be comparable across tumor samples (i.e., normalization). Additionally, in the case of TMAs, approaches to account for potential biases due to "holes" or regions in the image where cells were not able to be measured should be used, including permutation or Monte-Carlo methods. The mean of the empirical distribution computed under the assumption of spatial randomness can be used to compare to the observed value, producing an estimate of the degree of spatial clustering.

Lastly, further research is needed to develop methods and approaches that can combine information on the spatial TIME generated by mIF methods with other complementary methods used for studying the TIME (i.e., single-cell RNA-seq [22,23,145]; spatial transcriptomics [146–148]). Additionally, approaches that are able to determine tumor immune subtypes (i.e., "cold" vs. "hot" tumors, intratumoral immune states [149]) or "immunoscores" [11,13], using both abundance and spatial architecture of the TIME, are needed. Particularly these new approaches should accommodate for the zero-inflated/over-dispersed nature of the measurements within a repeated measurements analysis framework.

In conclusion, mIF technology has enabled the TIME to be studied at an increased scale, in terms of both the number of markers and the number of samples in a cost-effective manner. Another benefit of mIF technology is the ability to measure not only the abundance but also the spatial location of multiple cells types within a tissue sample simultaneously. This allows for the study of the co-localization and co-clustering of different types of immune markers simultaneously within both the tumor and stroma compartments. The development of methods and approaches that can deeply characterize the spatial heterogeneity of the TIME, classify immuno-phenotypes and identify immuno-spatial patterns could inform novel immunotherapy treatment approaches, as well as immunological biomarkers that can be used to predict immunotherapy response.

Author Contributions: Conceptualization, B.L.F.; statistical analysis, C.M.W., O.E.O., M.K.T.; resources, S.S.T., J.M.S., L.C.P.; data curation, M.K.T., O.E.O., J.N., C.M.S.; writing—original draft preparation, B.L.F., C.M.W., O.E.O., M.K.T., J.N., C.M.S.; writing—review and editing, B.L.F., C.M.W., J.M.S., L.C.P., S.S.T.; visualization, C.M.W., O.E.O. All authors have read and agreed to the published version of the manuscript.

Funding: The Nurses' Health Study and Nurses' Health Study II are supported by grants from the National Institutes of Health (UM1 CA186107, P01 CA87969, U01 CA176726). The funding for the mIF data generated on the AACES cohort was provided by K99/R00 CA218681. The AACES is supported by R01 CA237318 and R01 CA142081.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Boards of the various institutions. The NHS I and NHS II study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health (Advarra, Pro00022823 23 February 2021), and those of participating registries as required. The AACES study protocol was approved by the institutional review boards of each of the participating sites in AACES (protocol number 18018).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this review are available on request from the corresponding author.

Acknowledgments: The authors would like to acknowledge the Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, as the home of the Nurses' Health Study, which provided the data and images presented in this manuscript. We would like to thank the following state cancer registries for their help: A.L., A.Z., A.R., C.A., C.O., C.T., D.E., F.L., G.A., I.D., I.L., I.N., I.A., K.Y., L.A., M.E., M.D., M.A., M.I., N.E., N.H., N.J., N.Y., N.C., N.D., O.H., O.K., O.R., P.A., R.I., S.C., T.N., T.X., V.A., W.A., W.Y. We would also like to acknowledge and thank the AACES investigators, Anthony Alberg, Elisa Bandera, Jill Barnholtz-Sloan, Melissa Bondy, Michele Cote, Ellen Funkhouser, Patricia Moorman, Edward Peters, Ann Schwartz, and Paul Terry, for their contributions to the AACES. Some of the figures presented in this review paper were from data generated from the AACES study.

Conflicts of Interest: None of the authors have conflict of interest or financial disclosures to declare.

References

- 1. Fridman, W.H.; Pagès, F.; Sautès-Fridman, C.; Galon, J. The immune contexture in human tumours: Impact on clinical outcome. *Nat. Rev. Cancer* **2012**, *12*, 298–306. [CrossRef]
- 2. Havel, J.J.; Chowell, D.; Chan, T.A. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nat. Rev. Cancer* **2019**, *19*, 133–150. [CrossRef] [PubMed]
- 3. Couzin-Frankel, J. Breakthrough of the year 2013. Cancer immunotherapy. Science 2013, 342, 1432–1433. [CrossRef] [PubMed]
- 4. Ribas, A.; Wolchok, J.D. Cancer immunotherapy using checkpoint blockade. Science 2018, 359, 1350–1355. [CrossRef]
- 5. Drake, C.G.; Lipson, E.J.; Brahmer, J.R. Breathing new life into immunotherapy: Review of melanoma, lung and kidney cancer. *Nat. Rev. Clin. Oncol.* **2014**, *11*, 24–37. [CrossRef]
- 6. Menon, S.; Shin, S.; Dy, G. Advances in Cancer Immunotherapy in Solid Tumors. Cancers 2016, 8, 106. [CrossRef] [PubMed]
- Emens, L.A.; Ascierto, P.A.; Darcy, P.K.; Demaria, S.; Eggermont, A.M.M.; Redmond, S.L.; Seliger, B.; Marincola, F.M. Cancer immunotherapy: Opportunities and challenges in the rapidly evolving clinical landscape. *Eur. J. Cancer* 2017, *81*, 116–129. [CrossRef]
- 8. Fridman, W.H.; Zitvogel, L.; Suantès-Fridman, C.; Kroemer, G. The immune contexture in cancer prognosis and treatment. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 717–734. [CrossRef] [PubMed]
- 9. Gooden, M.J.; de Bock, G.H.; Leffers, N.; Daeman, T.; Nijman, H.W. The prognostic influence of tumour-infiltrating lymphocytes in cancer: A systematic review with meta-analysis. *Br. J. Cancer* **2011**, *105*, 93–103. [CrossRef]
- Galon, J.; Costes, A.; Sanchez-Cabo, F.; Kirilovsky, A.; Mlecnik, B.; Lagorce-Pagès, C.; Tosolini, M.; Camus, M.; Berger, A.; Wind, P.; et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* 2006, 313, 1960–1964. [CrossRef]
- Pagès, F.; Mlecnik, B.; Marliot, F.; Bindea, G.; Ou, F.; Bifulco, C.; Lugli, A.; Zlobec, I.; Rau, T.T.; Berger, M.D. International validation of the consensus Immunoscore for the classification of colon cancer: A prognostic and accuracy study. *Lancet* 2018, 391, 2128–2139. [CrossRef]
- 12. Galon, J.; Fridman, W.H.; Pages, F. The adaptive immunologic microenvironment in colorectal cancer: A novel perspective. *Cancer Res.* 2007, *67*, 1883–1886. [CrossRef] [PubMed]
- Angell, H.K.; Bruni, D.; Barrett, J.C.; Herbst, R.; Galon, J. The Immunoscore: Colon Cancer and Beyond. *Clin. Cancer Res.* 2020, 26, 332–339.
 [CrossRef]
- 14. Galon, J.; Lugli, A.; Bifulco, C.; Pagès, F.; Masucci, G.; Marincola, F.M.; Ascierto, P.A. World-Wide Immunoscore Task Force: Meeting report from the "Melanoma Bridge", Napoli, November 30th–December 3rd, 2016. J. Transl. Med. 2017, 15, 212. [CrossRef]
- 15. Galon, J.; Pagès, F.; Marincola, F.M.; Thurin, M.; Trinchieri, G.; Fox, B.A.; Gajewski, T.F.; Ascierto, P.A. The immune score as a new possible approach for the classification of cancer. *J. Transl. Med.* **2012**, *10*, 1. [CrossRef] [PubMed]
- 16. Galon, J.; Pagès, F.; Marincola, F.M.; Angell, H.; Thurin, M.; Lugli, A.; Zlobec, I.; Berger, A.; Bifulco, C.; Botti, G.; et al. Cancer classification using the Immunoscore: A worldwide task force. *J. Transl. Med.* **2012**, *10*, 205. [CrossRef]
- Yeong, J.; Tan, T.; Chow, Z.L.; Cheng, Q.; Lee, B.; Seet, A.; Lim, J.X.; Lim, J.C.T.; Ong, C.C.H.; Thike, A.A.; et al. Multiplex immunohistochemistry/immunofluorescence (mIHC/IF) for PD-L1 testing in triple-negative breast cancer: A translational assay compared with conventional IHC. J. Clin. Pathol. 2020, 73, 557–562. [CrossRef]
- 18. Newman, A.M.; Liu, C.L.; Green, M.R.; Gentles, A.J.; Feng, W.; Xu, Y.; Hoang, C.D.; Diehn, M.; Alizadeh, A.A. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **2015**, *12*, 453–457. [CrossRef]
- 19. Aran, D.; Hu, Z.; Butte, A.J. xCell: Digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 2017, *18*, 220. [CrossRef]
- 20. de Kanter, J.K.; Lijnzaad, P.; Candelli, T.; Margaritis, T.; Holstege, F.C.P. CHETAH: A selective, hierarchical cell type identification method for single-cell RNA sequencing. *Nucleic Acids Res.* **2019**, *47*, e95. [CrossRef]
- Wilson, C.M.; Fridley, B.L.; Conejo-Garcia, J.R.; Wang, X.; Yu, X. Wide and deep learning for automatic cell type identification. *Comput. Struct. Biotechnol. J.* 2021, 19, 1052–1062. [CrossRef]
- 22. Zheng, G.X.; Terry, J.M.; Belgrader, P.; Ryvkin, P.; Bent, Z.W.; Wilson, R.; Zirlado, S.B.; Wheeler, T.B.; McDermott, G.P.; Zhu, J.; et al. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **2017**, *8*, 14049. [CrossRef]

- 23. Yu, X.; Abbas-Aghababazadeh, F.; Chen, Y.A.; Fridley, B.L. Statistical and Bioinformatics Analysis of Data from Bulk and Single-Cell RNA Sequencing Experiments. *Methods Mol. Biol.* 2021, 2194, 143–175. [PubMed]
- 24. Young, Y.K.; Bolt, A.M.; Ahn, R.; Mann, K.K. Analyzing the Tumor Microenvironment by Flow Cytometry. *Methods Mol. Biol.* **2016**, *1458*, 95–110.
- Carvajal-Hausdorf, D.E.; Patsenker, J.; Stanton, K.P.; Villarroel-Espindola, F.; Esch, A.; Montgomery, R.R.; Psyrri, A.; Kalogeros, K.T.; Kotoula, V.; Foutzilas, G.; et al. Multiplexed (18-Plex) Measurement of Signaling Targets and Cytotoxic T Cells in Trastuzumab-Treated Patients using Imaging Mass Cytometry. *Clin. Cancer Res.* 2019, *25*, 3054–3062. [CrossRef]
- Giesen, C.; Wang, H.A.; Schapiro, D.; Zivanovic, N.; Jacobs, A.; Hattendorf, B.; Schuffler, P.J.; Grolimund, D.; Buhmann, J.M.; Brandt, S.; et al. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* 2014, 11, 417–422. [CrossRef]
- Martinez-Morilla, S.; Villarroel-Espindola, F.; Wong, P.F.; Toki, M.I.; Aung, T.N.; Pelekanou, V.; Bourke-Martin, B.; Schalper, K.A.; Kluger, H.M.; Rimm, D.L. Biomarker Discovery in Patients with Immunotherapy-Treated Melanoma with Imaging Mass Cytometry. *Clin. Cancer Res.* 2021, 27, 1987–1996. [CrossRef] [PubMed]
- 28. Baharlou, H.; Canete, N.P.; Cunningham, A.L.; Harman, A.N.; Patrick, E. Mass Cytometry Imaging for the Study of Human Diseases-Applications and Data Analysis Strategies. *Front Immunol.* **2019**, *10*, 2657. [CrossRef] [PubMed]
- Angelo, M.; Bendall, S.C.; Finck, R.; Hale, M.B.; Hitzman, C.; Borowsky, A.D.; Levenson, R.M.; Lowe, J.B.; Liu, S.D.; Zhao, S.; et al. Multiplexed ion beam imaging of human breast tumors. *Nat. Med.* 2014, 20, 436–442. [CrossRef]
- Liu, J.T.; Loewke, N.O.; Mandella, M.J.; Levenson, R.M.; Crawford, J.M.; Contag, C.H. Point-of-care pathology with miniature microscopes. *Anal. Cell Pathol.* 2011, 34, 81–98. [CrossRef]
- Sheffield, B.S. Immunohistochemistry as a Practical Tool in Molecular Pathology. Arch. Pathol. Lab. Med. 2016, 140, 766–769. [CrossRef] [PubMed]
- 32. Jones, M.L. Histotechnology a Self Instructional Text, 5th ed.; American Society of Clinical Oncology: Alexandria, VA, USA, 2020.
- 33. Prophet, E.B.; Armed, P. Forces Institute of, Laboratory Methods in Histotechnology; American Registry of Pathology: Washington, DC, USA, 1992.
- 34. Gnjatic, S.; Bronte, V.; Brunet, L.R.; Butler, M.O.; Disis, M.L.; Galon, J.; Hakansson, L.G.; Hanks, B.A.; Karanikas, V.; Khleif, S.N.; et al. Identifying baseline immune-related biomarkers to predict clinical outcome of immunotherapy. *J. Immunother. Cancer* **2017**, *5*, 44. [CrossRef]
- Herbst, R.S.; Soria, J.C.; Kowanetz, M.; Fine, G.D.; Hamid, O.; Gordon, M.S.; Sosman, J.A.; McDermott, D.F.; Powderly, J.D.; Gettinger, S.N.; et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* 2014, 515, 563–567. [CrossRef]
- 36. Hedvat, C.V. Digital microscopy: Past, present, and future. Arch. Pathol. Lab. Med. 2010, 134, 1666–1670. [CrossRef] [PubMed]
- 37. Taube, J.M.; Akturk, G.; Angelo, M.; Engle, E.L.; Gnjatic, S.; Greenbaum, S.; Greenwald, N.F.; Hedvat, C.V.; Hollmann, T.J.; Juco, J.; et al. The Society for Immunotherapy of Cancer statement on best practices for multiplex immunohistochemistry (IHC) and immunofluorescence (IF) staining and validation. J. Immunother. Cancer 2020, 8, 1. [CrossRef] [PubMed]
- Tsakiroglou, A.M.; Fergie, M.; Oguejiofor, K.; Linton, K.; Thomson, D.; Stern, P.L.; Astley, S.; Byers, R.; West, C.M.L. Spatial proximity between T and PD-L1 expressing cells as a prognostic biomarker for oropharyngeal squamous cell carcinoma. *Br. J. Cancer* 2020, *122*, 539–544. [CrossRef] [PubMed]
- Huang, Y.K.; Wang, M.; Sun, Y.; Di Costanzo, N.; Mitchell, C.; Achuthan, A.; Hamilton, J.A.; Busuttil, R.A.; Boussioutas, A. Macrophage spatial heterogeneity in gastric cancer defined by multiplex immunohistochemistry. *Nat. Commun.* 2019, 10, 3928. [CrossRef] [PubMed]
- Vayrynen, S.A.; Zhang, J.; Yuan, C.; Vayrynen, J.P.; Dias Costa, A.; Williams, H.; Morales-Oyarvide, V.; Lau, M.C.; Rubinson, D.A.; Dunne, R.F.; et al. Composition, Spatial Characteristics, and Prognostic Significance of Myeloid Cell Infiltration in Pancreatic Cancer. *Clin. Cancer Res.* 2021, 27, 1069–1081. [CrossRef]
- 41. Belanger, C.F.; Hennekens, C.H.; Rosner, B.; Speizer, F.E. The nurses' health study. Am. J. Nurs. 1978, 78, 1039–1040. [CrossRef]
- 42. Birmann, B.M.; Barnard, M.E.; Bertrand, K.A.; Bao, Y.; Crous-Bou, M.; Wolpin, B.M.; De Vivo, I.; Tworoger, S.S. Nurses' Health Study Contributions on the Epidemiology of Less Common Cancers: Endometrial, Ovarian, Pancreatic, and Hematologic. *Am. J. Public Health* **2016**, 106, 1608–1615. [CrossRef]
- 43. Boutot, M.E.; Purdue-Smithe, A.; Whitcomb, B.W.; Szegda, K.L.; Manson, J.E.; Hankinson, S.E.; Rosner, B.A.; Bertone-Johnson, E.R. Dietary Protein Intake and Early Menopause in the Nurses' Health Study II. *Am. J. Epidemiol.* **2018**, *187*, 270–277. [CrossRef]
- 44. Schildkraut, J.M.; Alberg, A.J.; Bandera, E.V.; Barnholtz-Sloan, J.; Bondy, M.; Cote, M.L.; Funkhouser, E.; Peters, E.; Schwartz, A.G.; Terry, P.; et al. A multi-center population-based case-control study of ovarian cancer in African-American women: The African American Cancer Epidemiology Study (AACES). *BMC Cancer* **2014**, *14*, 688. [CrossRef]
- Biswas, S.; Mandal, G.; Payne, K.K.; Anadon, C.M.; Gatenbee, C.D.; Chaurio, R.A.; Costich, T.L.; Moran, C.; Harro, C.M.; Rigolizzo, K.E.; et al. IgA transcytosis and antigen recognition govern ovarian cancer immunity. *Nature* 2021, 591, 464–470. [CrossRef] [PubMed]
- 46. Hajiran, A.; Chakiryan, N.; Aydin, A.M.; Zemp, L.; Nguyen, J.; Laborde, J.M.; Chahoud, J.; Spiess, P.E.; Zaman, S.; Falasiri, S.; et al. Reconnaissance of tumor immune microenvironment spatial heterogeneity in metastatic renal cell carcinoma and correlation with immunotherapy response. *Clin. Exp. Immunol.* **2021**, 204, 96–106. [CrossRef]

- Kamal, Y.; Dwan, D.; Hoehn, H.J.; Sanz-Pamplona, R.; Alonso, M.H.; Moreno, V.; Cheng, C.; Schell, M.J.; Kim, Y.; Felder, S.I. Tumor immune infiltration estimated from gene expression profiles predicts colorectal cancer relapse. *OncoImmunology* 2021, 10, 1862529. [CrossRef] [PubMed]
- 48. Akoya Biosciences. Opal Mulitplex IHC Assay Development Guide; Akoya Biosciences: Marlborough, MA, USA, 2019.
- Lim, J.C.T.; Yeong, J.P.S.; Lim, C.J.; Ong, C.C.H.; Wong, S.C.; Chew, V.S.P.; Ahmed, S.S.; Tan, P.H.; Iqbal, J. An automated staining protocol for seven-colour immunofluorescence of human tissue sections for diagnostic and prognostic use. *Pathology* 2018, 50, 333–341. [CrossRef]
- 50. Garini, Y.; Young, I.T.; McNamara, G. Spectral imaging: Principles and applications. Cytometry A 2006, 69, 735–747. [CrossRef]
- Abel, E.J.; Bauman, T.M.; Weiker, M.; Shi, F.; Downs, T.M.; Jarrard, D.F.; Huang, W. Analysis and validation of tissue biomarkers for renal cell carcinoma using automated high-throughput evaluation of protein expression. *Hum. Pathol.* 2014, 45, 1092–1099. [CrossRef]
- 52. Ghaznavi, F.; Evans, A.; Madabhushi, A.; Feldman, M. Digital imaging in pathology: Whole-slide imaging and beyond. *Annu. Rev. Pathol.* **2013**, *8*, 331–359. [CrossRef] [PubMed]
- Parra, E.R.; Uraoka, N.; Jiang, M.; Cook, P.; Gibbons, D.; Forget, M.-A.; Bernatchez, C.; Haymaker, C.; Wistuba, I.I.; Rodriguez-Canales, J. Validation of multiplex immunofluorescence panels using multispectral microscopy for immune-profiling of formalinfixed and paraffin-embedded human tumor tissues. *Sci. Rep.* 2017, *7*, 13380. [CrossRef]
- 54. Acs, B.; Pelekanou, V.; Bai, Y.; Martinez-Morilla, S.; Toki, M.; Leung, S.C.Y.; Nielsen, T.O.; Rimm, D.L. Ki67 reproducibility using digital image analysis: An inter-platform and inter-operator study. *Lab. Investig.* **2019**, *99*, 107–117. [CrossRef]
- 55. Horai, Y.; Mizukawa, M.; Nishina, H.; Nishikawa, S.; Ono, Y.; Takemoto, K.; Baba, N. Quantification of histopathological findings using a novel image analysis platform. *J. Toxicol. Pathol.* **2019**, *32*, 319–327. [CrossRef] [PubMed]
- 56. Shakya, R.; Nguyen, T.H.; Waterhouse, N.; Khanna, R. Immune contexture analysis in immuno-oncology: Applications and challenges of multiplex fluorescent immunohistochemistry. *Clin. Transl. Immunol.* **2020**, *9*, e1183. [CrossRef] [PubMed]
- Stack, E.C.; Wang, C.; Roman, K.A.; Hoyt, C.C. Multiplexed immunohistochemistry, imaging, and quantitation: A review, with an assessment of Tyramide signal amplification, multispectral imaging and multiplex analysis. *Methods* 2014, 70, 46–58. [CrossRef] [PubMed]
- Tan, W.C.C.; Nerurkar, S.N.; Cai, H.Y.; Ng, H.H.M.; Wu, D.; Wee, Y.T.F.; Lim, J.C.T.; Yeong, J.; Lim, T.K.H. Overview of multiplex immunohistochemistry/immunofluorescence techniques in the era of cancer immunotherapy. *Cancer Commun.* 2020, 40, 135–153. [CrossRef]
- Gorris, M.A.J.; Halilovic, A.; Rabold, K.; van Duffelen, A.; Wickramasinghe, I.N.; Verweij, D.; Wortel, I.M.N.; Textor, J.C.; de Vries, I.J.M.; Figdor, C.G. Eight-Color Multiplex Immunohistochemistry for Simultaneous Detection of Multiple Immune Checkpoint Molecules within the Tumor Microenvironment. *J. Immunol.* 2018, 200, 347–354. [CrossRef]
- 60. Mezheyeuski, A.; Bergsland, C.H.; Backman, M.; Djureinovic, D.; Sjoblom, T.; Bruun, J.; Micke, P. Multispectral imaging for quantitative and compartment-specific immune infiltrates reveals distinct immune profiles that classify lung cancer patients. *J. Pathol.* **2018**, *244*, 421–431. [CrossRef]
- 61. Mori, H.; Bolen, J.; Schuetter, L.; Massion, P.; Hoyt, C.C.; VandenBerg, S.; Esserman, L.; Borowsky, A.D.; Campbell, M.J. Characterizing the Tumor Immune Microenvironment with Tyramide-Based Multiplex Immunofluorescence. *J. Mammary Gland. Biol. Neoplasia* **2020**, *25*, 417–432. [CrossRef]
- Amancio, D.R.; Comin, C.H.; Casanova, D.; Travieso, G.; Bruno, O.M.; Rodrigues, F.A.; Costa Lda, F. A systematic comparison of supervised classifiers. *PLoS ONE* 2014, 9, e94137. [CrossRef] [PubMed]
- 63. Breiman, L. Random Forests. Mach. Learn. 2001, 45, 5–32. [CrossRef]
- 64. Thurin, M.; Cesano, A.; Marincola, F. *Biomarkers for Immunotherapy of Cancer. Methods in Molecular Biology*; Thurin, M., Cesano, A., Marincola, F., Eds.; Springer: New York, NY, USA, 2020.
- 65. Viratham Pulsawatdi, A.; Craig, S.G.; Bingham, V.; McCombe, K.; Humphries, M.P.; Senevirathne, S.; Richman, S.D.; Quirke, P.; Campo, L.; Domingo, E.; et al. A robust multiplex immunofluorescence and digital pathology workflow for the characterisation of the tumour immune microenvironment. *Mol. Oncol.* **2020**, *14*, 2384–2402. [CrossRef]
- Blessin, N.C.; Simon, R.; Kluth, M.; Fischer, K.; Hube-Magg, C.; Li, W.; Makrypidi-Fraune, G.; Wellge, B.; Mandelkow, T.; Debatin, N.F.; et al. Patterns of TIGIT Expression in Lymphatic Tissue, Inflammation, and Cancer. *Dis. Markers* 2019, 2019, 5160565. [CrossRef] [PubMed]
- 67. Govek, K.W.; Troisi, E.C.; Miao, Z.; Aubin, R.G.; Woodhouse, S.; Camara, P.G. Single-cell transcriptomic analysis of mIHC images via antigen mapping. *Sci. Adv.* 2021, *7*, eabc5464. [CrossRef]
- 68. Tworoger, S.S.; Hankinson, S.E. Use of biomarkers in epidemiologic studies: Minimizing the influence of measurement error in the study design and analysis. *Cancer Causes Control.* **2006**, *17*, 889–899. [CrossRef] [PubMed]
- Parra, E.R.; Jiang, M.; Solis, L.; Mino, B.; Laberiano, C.; Hernandez, S.; Gite, S.; Verma, A.; Tetzlaff, M.; Haymaker, C.; et al. Procedural Requirements and Recommendations for Multiplex Immunofluorescence Tyramide Signal Amplification Assays to Support Translational Oncology Studies. *Cancers* 2020, *12*, 255. [CrossRef]
- 70. Lee, C.-W.; Ren, Y.J.; Marella, M.; Wang, M.; Hartke, J.; Couto, S.S. Multiplex immunofluorescence staining and image analysis assay for diffuse large B cell lymphoma. *J. Immunol. Methods* **2020**, *478*, 112714. [CrossRef]
- 71. Diem, K.; Magaret, A.; Klock, A.; Jin, L.; Zhu, J.; Corey, L. Image analysis for accurately counting CD4+ and CD8+ T cells in human tissue. *J. Virol. Methods.* 2015, 222, 117–121. [CrossRef]

- 72. Sanchez, K.; Kim, I.; Chun, B.; Pucilowska, J.; Redmond, W.L.; Urba, W.J.; Martel, M.; Wu, Y.; Campbell, M.; Sun, Z.; et al. Multiplex immunofluorescence to measure dynamic changes in tumor-infiltrating lymphocytes and PD-L1 in early-stage breast cancer. *Breast Cancer Res.* **2021**, *23*, 2. [CrossRef]
- 73. McCullagh, P.; Nelder, J.A. Generalized Linear Models; Chapman & Hall/CRC: Boca Raton, FL, USA, 1999; p. 511.
- 74. Agresti, A. *Categorical Data Analysis*, 2nd ed.; Wiley Series in Probability and Statistics; ohn Wiley & Sons, Inc.: Hoboken, NJ, USA, 2002; p. 710.
- 75. Genser, B.; Cooper, P.J.; Yazdanbakhsh, M.; Barreto, M.L.; Rodrigues, L.C. A guide to modern statistical analysis of immunological data. *BMC Immunol.* 2007, *8*, 27. [CrossRef] [PubMed]
- 76. Ramsey, F.L.; Schafer, D.W. *The Statistical Sleuth: A Course in Methods of Data Analysis*, 3rd ed.; Brooks/Cole, Cengage Learning: Boston, MA, USA, 2013.
- Hammond, M.E.H.; Hayes, D.F.; Dowsett, M.; Allred, D.C.; Hagerty, K.L.; Badve, S.; Fitzgibbons, P.L.; Francis, G.; Goldstein, N.S.; Hayes, M.; et al. American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations for Immunohistochemical Testing of Estrogen and Progesterone Receptors in Breast Cancer. J. Clin. Oncol. 2010, 28, 2784–2795. [CrossRef]
- Borghaei, H.; Paz-Ares, L.; Horn, L.; Spigel, D.R.; Steins, M.; Ready, N.E.; Chow, L.Q.; Vokes, E.E.; Felip, E.; Holgado, E.; et al. Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. N. Engl. J. Med. 2015, 373, 1627–1639. [CrossRef] [PubMed]
- Dong, H.; Strome, S.E.; Salomao, D.R.; Tamura, H.; Hirano, F.; Flies, D.B.; Roche, P.C.; Lu, J.; Zhu, G.; Tamada, K.; et al. Tumor-associated B7-H1 promotes T-cell apoptosis: A potential mechanism of immune evasion. *Nat. Med.* 2002, *8*, 793–800. [CrossRef]
- Patel, S.P.; Kurzrock, R. PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. *Mol. Cancer Ther.* 2015, 14, 847. [CrossRef]
- 81. Ilie, M.; Hofman, V.; Dietel, M.; Soria, J.C.; Hofman, P. Assessment of the PD-L1 status by immunohistochemistry: Challenges and perspectives for therapeutic strategies in lung cancer patients. *Virchows Arch.* **2016**, *468*, 511–525. [CrossRef] [PubMed]
- 82. Ribas, A.; Hu-Lieskovan, S. What does PD-L1 positive or negative mean? J. Exp. Med. 2016, 213, 2835–2840. [CrossRef]
- 83. Bouwmeester, W.; Zuithoff, N.P.; Mallett, S.; Geerlings, M.I.; Vergouwe, Y.; Steyerberg, E.W.; Altman, D.G.; Moons, K.G. Reporting and methods in clinical prediction research: A systematic review. *PLoS Med.* **2012**, *9*, 1–12. [CrossRef]
- 84. Mabikwa, O.V.; Greenwood, D.C.; Baxter, P.D.; Fleming, S.J. Assessing the reporting of categorised quantitative variables in observational epidemiological studies. *BMC Health Serv. Res.* 2017, 17, 201.
- 85. Altman, D.G.; Lausen, B.; Sauerbrei, W.; Schumacher, M. Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. J. Natl. Cancer Inst. 1994, 86, 829–835. [CrossRef]
- Wilson, C.; Thapa, R.; Creed, J.; Nguyen, J.; Segura, C.M.; Gerke, T.; Schildkraut, K.; Peres, L.; Fridley, B.L. Statistical framework for studying the spatial architecture of the tumor immune microenvironment. *medRxiv* 2021. [CrossRef]
- 87. Hall, D.B. Zero-inflated Poisson and binomial regression with random effects: A case study. *Biometrics* **2000**, *56*, 1030–1039. [CrossRef]
- 88. Lee, A.H.; Lee, A.H.; Wang, K.; Scott, J.A.; Yau, K.K.; McLachlan, G.J. Multi-level zero-inflated poisson regression modelling of correlated count data with excess zeros. *Stat. Methods Med. Res.* **2006**, *15*, 47–61. [CrossRef]
- 89. Jiang, S.; Xiao, G.; Koh, A.Y.; Kim, J.; Li, Q.; Zhan, X. A Bayesian zero-inflated negative binomial regression model for the integrative analysis of microbiome data. *Biostatistics* **2019**. [CrossRef]
- 90. Zhang, X.; Yi, N. Fast zero-inflated negative binomial mixed modeling approach for analyzing longitudinal metagenomics data. *Bioinformatics* **2020**, *36*, 2345–2351. [CrossRef] [PubMed]
- 91. Hu, T.; Gallins, P.; Zhou, Y.H. A Zero-inflated Beta-binomial Model for Microbiome Data Analysis. *Stat. Int. Stat. Inst.* 2018, 7, e185. [CrossRef] [PubMed]
- 92. Yau, K.K.; Lee, A.H. Zero-inflated Poisson regression with random effects to evaluate an occupational injury prevention programme. *Stat. Med.* 2001, 20, 2907–2920. [CrossRef]
- 93. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef]
- 94. Robinson, M.D.; Smyth, G.K. Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics* 2007, 23, 2881–2887. [CrossRef]
- 95. Robinson, M.D.; Smyth, G.K. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* **2008**, *9*, 321–332. [CrossRef] [PubMed]
- 96. J-Express 2012, Visualization and Analysis of Microarray Data. pp. 161–177. Available online: https://mybiosoftware.com/jexpress-2009-analysis-visualization-microarray-data.html (accessed on 20 May 2021).
- Jeong, H.H.; Kim, S.Y.; Rousseaux, M.W.C.; Zoghbi, H.Y.; Liu, Z. Beta-binomial modeling of CRISPR pooled screen data identifies target genes with greater sensitivity and fewer false negatives. *Genome Res.* 2019, 29, 999–1008. [CrossRef]
- 98. Kim, J.; Lee, J.H. The Validation of a Beta-Binomial Model for Overdispersed Binomial Data. *Commun. Stat. Simul. Comput.* 2017, 46, 807–814. [CrossRef]
- 99. Martin, B.D.; Witten, D.; Willis, A.D. Modeling Microbial Abundances and Dysbiosis with Beta-Binomial Regression. *Ann. Appl. Stat.* **2020**, *14*, 94–115. [CrossRef]

- 100. Najera-Zuloaga, J.; Lee, D.J.; Arostegui, I. Comparison of beta-binomial regression model approaches to analyze health-related quality of life data. *Stat. Methods Med. Res.* **2018**, *27*, 2989–3009. [CrossRef]
- Jakaitiene, A.; Avino, M.; Guarracino, M.R. Beta-Binomial Model for the Detection of Rare Mutations in Pooled Next-Generation Sequencing Experiments. J. Comput. Biol. 2017, 24, 357–367. [CrossRef]
- 102. Congdon, P. Bayesian Statistical Modelling. Wiley Series in Probability and Statistics; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2007; p. 552.
- 103. McCulloch, C.E.; Searle, S.R. Generalized, Linear, and Mixed Models. Wiley Series in Probability and Statistics Texts, References, and Pocketbooks Section; John Wiley & Sons, Inc.: New York, NY, USA, 2001; p. 325.
- 104. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B* **1977**, *39*, 1–38.
- 105. Ypma, T.J. Historical Development of the Newton-Raphson Method. Soc. Ind. Appl. Math. Rev. 1995, 37, 531–551. [CrossRef]
- 106. Geman, S.; Geman, D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **1984**, *6*, 721–741. [CrossRef]
- 107. Gilks, W.R.; Richardson, S.; Spiegelhalter, D.J. *Markov Chain Monte Carlo in Practice*; Chapman & Hall/CRC: Boca Raton, FL, USA, 1996; p. 486.
- 108. Sainani, K. The importance of accounting for correlated observations. Pm R. 2010, 2, 858–861. [CrossRef] [PubMed]
- Schober, P.; Vetter, T.R. Repeated Measures Designs and Analysis of Longitudinal Data: If at First You Do Not Succeed-Try, Try Again. Anesth Analg. 2018, 127, 569–575. [CrossRef]
- 110. Magurran, A.E. Biological Diversity. Curr. Biol. 2005, 15, R116–R118. [CrossRef] [PubMed]
- 111. Horn, H.S. Measurement of overlap in comparative ecological studies. Am. Nat. 1966, 100, 419–424. [CrossRef]
- 112. Wolda, H. Similarity indices, sample size and diversity. *Oecologia* **1981**, *50*, 296–302. [CrossRef] [PubMed]
- 113. Duncan, O.D.; Duncan, B. A Methodological Analysis of Segregation Indexes. Am. Sociol. Rev. 1955, 20, 210–217. [CrossRef]
- Yao, J.; Wong, D.; Bailey, N.; Minton, J. Spatial Segregation Measures: A Methodological Review: Spatial Segregation Measures. *Tijdschr. Econ. Soc. Geogr.* 2018, 110, 235–250. [CrossRef]
- 115. Ripley, B.D. The second-order analysis of stationary point processes. J. Appl. Probab. 1976, 13, 255–266. [CrossRef]
- 116. Baddeley, A.; Rubak, E.; Turner, R. Spatial Point Patterns: Methodology and Applications with R; CRC: Boca Raton, FL, USA, 2015; p. 810.
- 117. Besag, J. Comments on Ripley's paper. J. R. Stat. Soc. Ser. A 1977, 39, 193-195.
- 118. Marcon, E.; Puech, F.; Traissac, S. Characterizing the Relative Spatial Structure of Point Patterns. *Int. J. Ecol.* 2012, 2012, 619281. [CrossRef]
- 119. Bull, J.A.; Macklin, P.S.; Quaiser, T.; Braun, F.; Waters, S.L.; Pugh, C.W.; Byrne, H.M. Combining multiple spatial statistics enhances the description of immune cell localisation within tumours. *Sci. Rep.* **2020**, *10*, 18624. [CrossRef]
- Agnew, D.; Green, J.; Brown, T.M.; Simpson, M.J.; Binder, B.J. Distinguishing between mechanisms of cell aggregation using pair-correlation functions. J. Theor. Biol. 2014, 352, 16–23. [CrossRef]
- Rose, C.J.; Naidoo, K.; Clay, V.; Linton, K.; Radford, J.A.; Byers, R.J. A statistical framework for analyzing hypothesized interactions between cells imaged using multispectral microscopy and multiple immunohistochemical markers. J. Pathol. Inf. 2013, 4 (Suppl. 4). [CrossRef]
- 122. Baddeley, A.J.; Gill, R.D. *The Empty Space Hazard of a Spatial Pattern*; University of Western Australia. Department of Mathematics: Perth, Australia, 1994.
- 123. Baddeley, A.; Gill, R.D. Kaplan-Meier Estimators of Distance Distributions for Spatial Point Processes. *Ann. Stat.* **1997**, *25*, 263–292. [CrossRef]
- 124. Baddeley, A.J.; Kerscher, M.; Schladitz, K.; Scott, B.T. Estimating the J function without edge correction. *Stat. Neerl.* 2000, 54, 315–328. [CrossRef]
- 125. Costes, S.V.; Daelemans, D.; Cho, E.H.; Dobbin, Z.; Pavlakis, G.; Lockett, S. Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophys. J.* **2004**, *86*, 3993–4003. [CrossRef] [PubMed]
- 126. Yuan, Y. Spatial Heterogeneity in the Tumor Microenvironment. Cold Spring Harb. Perspect. Med. 2016, 6, a026583. [CrossRef]
- 127. Maley, C.C.; Koelble, K.; Natrajan, R.; Aktipis, A.; Yuan, Y. An ecological measure of immune-cancer colocalization as a prognostic factor for breast cancer. *Breast Cancer Res. BCR* 2015, 17, 131. [CrossRef] [PubMed]
- 128. Rempala, G.A.; Seweryn, M. Methods for diversity and overlap analysis in T-cell receptor populations. *J. Math. Biol.* **2013**, *67*, 1339–1368. [CrossRef]
- Roh, K.-H.; Lillemeier, B.F.; Wang, F.; Davis, M.M. The coreceptor CD4 is expressed in distinct nanoclusters and does not colocalize with T-cell receptor and active protein tyrosine kinase p56lck. *Proc. Natl. Acad. Sci. USA* 2015, 112, E1604. [CrossRef] [PubMed]
- 130. Kang, J.M. Voronoi Diagram. In Encyclopedia of GIS; Shekhar, S., Xiong, H., Eds.; Springer: Boston, MA, USA, 2008; pp. 1232–1235.
- Enfield, K.S.S.; Martin, S.D.; Marshall, E.A.; Kung, S.H.Y.; Gallagher, P.; Milne, K.; Chen, Z.; Nelson, B.H.; Lam, S.; English, J.C.; et al. Hyperspectral cell sociology reveals spatial tumor-immune cell interactions associated with lung cancer recurrence. J. Immunother. Cancer 2019, 7, 13. [CrossRef] [PubMed]
- 132. Gartrell, R.D.; Marks, D.K.; Hart, T.D.; Li, G.; Davari, D.R.; Wu, A.; Blake, Z.; Lu, Y.; Askin, K.N.; Monod, A.; et al. Quantitative Analysis of Immune Infiltrates in Primary Melanoma. *Cancer Immunol. Res.* **2018**, *6*, 481–493. [CrossRef]

- 133. Parra, E.R.; Zhai, J.; Tamegnon, A.; Zhou, N.; Pandurengan, R.K.; Barreto, C.; Jiang, M.; Rice, D.C.; Creasy, C.; Vaporciyan, A.A.; et al. Identification of distinct immune landscapes using an automated nine-color multiplex immunofluores-cence staining panel and image analysis in paraffin tumor tissues. *Sci. Rep.* 2021, *11*, 4530. [CrossRef]
- 134. Fassler, D.J.; Abousamra, S.; Gupta, R.; Chen, C.; Zhao, M.; Paredes, D.; Batool, S.A.; Knudsen, B.S.; Escobar-Hoyos, L.; Shroyer, K.R.; et al. Deep learning-based image analysis methods for brightfield-acquired multiplex immunohistochemistry images. *Diagn. Pathol.* 2020, 15, 100. [CrossRef] [PubMed]
- 135. Moore, M. Spatial Statistics: Methodological Aspects and Applications; Springer: New York, NY, USA, 2001; p. 282.
- 136. Cressie, N.A.C. *Statistics for Spatial Data. Revised Edition;* Wiley Series in Probability and Mathematical Statistics; John Wiley & Sons, Inc.: New York, NY, USA, 1993; p. 900.
- 137. Barua, S.; Fang, P.; Sharma, A.; Fujimoto, J.; Wistuba, I.; Rao, A.U.K.; Lin, S.H. Spatial interaction of tumor cells and regulatory T cells correlates with survival in non-small cell lung cancer. *Lung Cancer* **2018**, *117*, 73–79. [CrossRef]
- Binder, B.J.; Simpson, M.J. Quantifying spatial structure in experimental observations and agent-based simulations using pair-correlation functions. *Phys. Rev. E* 2013, *88*, 022705. [CrossRef]
- 139. Cressie, N.; Hoef, J.M.V. Spatial Statistical Analysis of Environmental and Ecological Data; Oxford University Press: Oxford, UK, 1993; pp. 404–409.
- 140. Gabriel, E.; Baddeley, A.; Rubak, E.; Turner, R. Spatial Point Patterns: Methodology and Applications with R. *Math. Geosci.* **2017**, *49*, 815–817. [CrossRef]
- 141. Ramsay, J.O.; Silverman, B.W. *Applied Functional Data Analysis: Methods and Case Studies*; Springer Series in Statistics; Springer: New York, NY, USA, 2002.
- Leek, J.T.; Scharpf, R.B.; Bravo, H.C.; Simcha, D.; Langmead, B.; Johnson, W.E.; Geman, D.; Baggerly, K.; Irizarry, R.A. Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat. Rev. Genet.* 2010, *11*, 733–739. [CrossRef] [PubMed]
- 143. Leek, J.T.; Storey, J.D. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet.* 2007, *3*, 1724–1735. [CrossRef] [PubMed]
- 144. Johnson, W.E.; Li, C.; Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **2007**, *8*, 118–127. [CrossRef]
- 145. Tirosh, I.; Izar, B.; Prakadan, S.M.; Wadsworth, M.H., 2nd; Treacy, D.; Trombetta, J.J.; Rotem, A.; Rodman, C.; Lian, C.; Murphy, G.; et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 2016, 352, 189–196. [CrossRef] [PubMed]
- Thrane, K.; Eriksson, H.; Maaskola, J.; Hansson, J.; Lundeberg, J. Spatially Resolved Transcriptomics Enables Dissection of Genetic Heterogeneity in Stage III Cutaneous Malignant Melanoma. *Cancer Res.* 2018, 78, 5970–5979. [CrossRef]
- 147. Stahl, P.L.; Salmen, F.; Vickovic, S.; Lundmark, A.; Navarro, J.F.; Magnusson, J.; Giacomello, S.; Asp, M.; Westholm, J.O.; Huss, M.; et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **2016**, *353*, 78–82. [CrossRef]
- 148. Berglund, E.; Maaskola, J.; Schultz, N.; Friedrich, S.; Marklund, M.; Bergenstrahle, J.; Tarish, F.; Tanoglidi, A.; Vickovic, S.; Larsson, L.; et al. Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity. *Nat. Commun.* 2018, 9, 2419. [CrossRef] [PubMed]
- 149. Thorsson, V.; Gibbs, D.L.; Brown, S.D.; Wolf, D.; Bortone, D.S.; Ou Yang, T.H.; Porta-Pardo, E.; Gao, G.F.; Plaisier, C.L.; Eddy, J.A.; et al. The Immune Landscape of Cancer. *Immunity* **2018**, *48*, 812–830. [CrossRef]