Estimation of Infiltrating Lymphocytes in CNS Tumors from DNA Methylation Profiles

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Zahnmedizin (Dr. med. dent.)

an der

Medizinischen Fakultät der Universität Hamburg

vorgelegt von

Sepehr Safaei

aus

Esfahan

2025

Betreuer:in / Gutachter:in der Dissertation: Prof. Dr. Ulrich Schüller

Gutachter:in der Dissertation: Prof. Dr. Julia Neumann

Vorsitz der Prüfungskommission: Prof. Dr. Julia Neumann

Mitglied der Prüfungskommission: PD Dr. Georg Rosenberger

Mitglied der Prüfungskommission: Prof. Dr. Michael Bockmayr

Datum der mündlichen Prüfung: 10.12.2025

Table of Contents

1. Presentation of the Publication

1.1. Introduction

Profiling tumor-infiltrating lymphocytes (TILs) delivers valuable diagnostic and prognostic information (Hanahan and Weinberg 2011, Quail and Joyce 2017). It can be done via different approaches. One of the classical methods is scoring HE-stained slides (Hendry et al. 2017), which can be further optimized by machine learning (Klauschen et al. 2018). For the cases, in which tumors are morphologically similar to immune cells, immunohistochemistry is more effective. For rare brain tumors with significant variability in cellular composition, both of these methods may be inadequate. In such cases, molecular analyses would be a better choice. A main challenge in this method is profiling immune cells from bulk molecular data using deconvolution algorithms. There are various molecular analysis methods. In our study, we used methylation analysis and took gene expression-based estimates and immunohistochemical counts of TIL as the benchmarks. We chose DNA methylation-based profiling because it is widely used in molecular analyses of brain tumors (Cavalli et al. 2017, Northcott et al. 2017, Pajtler et al. 2015, Torchia et al. 2016, Sturm et al. 2016) and routine clinical diagnostics (Capper et al. 2018, Jurmeister et al. 2019, Orozco et al. 2018, Moran et al. 2016).

There are several approaches for gene expression-based profiling of immune cells:

- In one method, genes specific to each immune cell type are identified, and the average of their expression values is considered as the abundance score of immune cells. (Becht et al. 2016, Danaher et al. 2017).
- Another approach is CIBERSORT (Newman et al. 2015), which uses a deconvolution algorithm based on support vector regression. The outcome of this method is estimated proportions for 22 types of immune cells. For each estimation, the algorithm gives a p-value as well, which assesses the reliability of the estimates (Gentles et al. 2015).

Similar to gene expression-based profiling, DNA methylation-based profiling of immune cells can be conducted in various ways:

- In MeTIL, CpG sites specific to each immune cell type are identified, and a non-negative least squares (NNLS) regression model is used to estimate the abundance of immune cells (Jeschke et al. 2017).
- Another approach is MethylCIBERSORT, which is an adaptation of CIBERSORT for methylation data (Chakravarthy et al. 2018). This method is based on tumor-specific methylation signatures, which can be challenging for rare brain tumors.

Figure 1 presents an overview of the aforementioned methods.

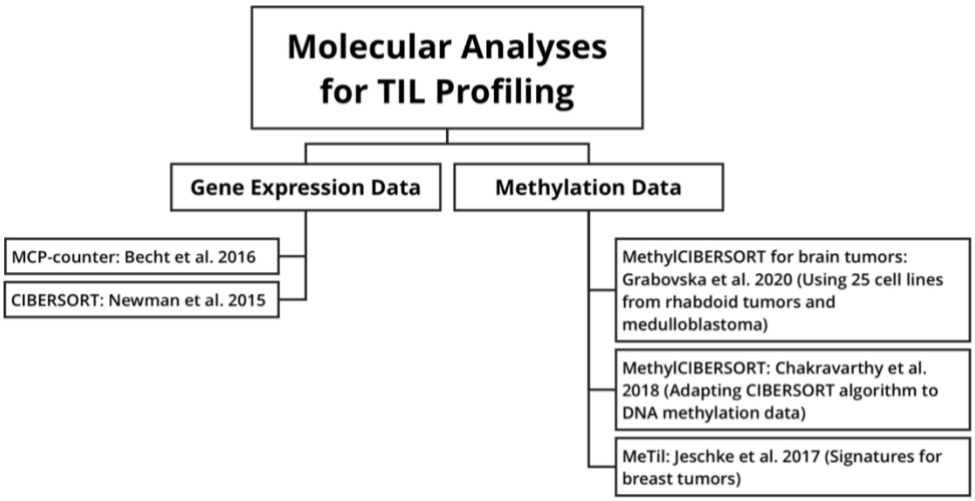


Figure 1. Overview of methods in molecular analysis

1.2. Materials and Methods

We employed the statistical programming language R version 3.6.0 (The R Core Team 2019) for data analysis. The following R packages were used: missMethyl (Phipson et al. 2016), minfi (Aryee et al. 2014), lumi (Du et al. 2008), ComplexHeatmap (Gu et al. 2016), MASS (Ripley et al. 2020), beeswarm (Eklund 2016), TCGAbiolinks (Colaprico et al. 2016), MethylCIBERSORT (Chakravarthy et al. 2018), FlowSorted.Blood.450k (Jaffe 2020), and survival (Therneau and Grambsch 2000).

1.2.1. Datasets and Preprocessing

The data used in this study can be categorized into three groups.

1. Training samples: based on which we obtained the methylation signatures and trained our algorithm.
2. Validation samples: based on which we tested our algorithm as well as other previously published methods and compared the results to validate our method's performance.
3. Clinical samples: to demonstrate the clinical application of our method.

All raw data were previously published and analyzed on the Illumina Infinium Methylation 450K BeadChip and are available from public data repositories. Table 1 (Safaei et al. 2021) outlines the quantities and origins of the used data.

Dataset Type	Cell Type	ID	Author	Number of Cases
Training Dataset	Brain Tumor (85 Entities)	GSE90496	Capper 2018	2706
	$CD4^+$	GSE59065	Tserel 2015	94
	$CD8^+$	GSE59065	Tserel 2015	94
	Different Cells in Blood	GSE35069	Reinius 2012	60
Validation Dataset	Medulloblastoma	GSE85212	Cavalli 2018	763 (763)*
	Ependymoma	GSE64415	Pajtler 2015	557 (129)*
	ATRT	27960086	Torchia 2017	162 (88)*
	Low Grade Glioma	TCGA-LGG		534 (532)*
	Glioblastoma	TCGA-GBM		155 (64)*

* Number of samples for which both gene expression and methylation data are available.

Table 1: Data sources and number of methylation data samples (Safaei et al. 2021)

Training Samples

To train our algorithm, we utilized three datasets.

1. As the reference for central nervous system (CNS) tumors, the training cohort of the Brain Tumor Classifier published by Capper et al. was used (Capper et al. 2018). Because differential methylation was conducted between bulk tumors and lymphocytes, samples from reference tissue and tumors with anticipated high lymphocyte infiltration were excluded. The excluded methylation classes are LYMPHO, PLASMA, CONTR INFLAM, CONTR REACT, MELAN, and MELCYT. This resulted in a dataset of 2706 samples from 85 diagnostic categories.
2. As the reference for T cells, methylation profiles of magnetically activated cell-sorted CD4+ and CD8+ T cells from 101 cases were obtained from Gene Expression Omnibus (GEO) (Tserel et al. 2015, Edgar et al. 2002). Since not all samples contained data for both CD4+ and CD8+ T cells, profiles from patients missing either CD4+ or CD8+ data were excluded, resulting in a dataset of 94 samples.
3. As the reference for peripheral blood cells, methylation profiles from the peripheral blood were obtained from Reinius et al., who conducted flow cytometric analyses on CD4+, CD8+ T cells, monocytes (CD14+), NK cells (CD56+), B cells (CD19+), as well as neutrophils, eosinophils, (mixed) granulocytes and (mixed) peripheral blood mononuclear cells (PBMCs) of six patients (Reinius et al. 2012). To obtain this dataset the R package flow.sorted.blood.450k. was used (Jaffe 2020).

Validation Samples

To validate our algorithm, we employed two types of datasets.

1. Samples for which methylation profiles as well as gene expression-based TIL scores were available or obtained. An overview of these datasets is presented in Table 1 (Safaei et al. 2021). Further detailed explanations follow below.
   - 763 methylation profiles from medulloblastoma (Cavalli et al. 2017) were employed. Expression data for all these samples were analyzed on the Affymetrix Human Gene 1.1 ST Array (Bockmayr et al. 2019).
   - Out of 557 methylation profiles of ependymoma (Pajtler et al. 2015), expression data for 129 samples were analyzed on the Affymetrix HG U133 Plus 2.0 microarray and preprocessed with the R package affy using the custom chip definition file hgu133plus2hsentrezgcdf (v19.0.0) (Gautier et al. 2004, Dai et al. 2005).
   - Out of 162 methylation profiles of atypical teratoid/rhabdoid tumor (ATRT) (Torchia et al. 2015, Torchia et al. 2016), expression data for 88 samples were analyzed on the Illumina HT12 gene expression array and preprocessed with the R package lumi (Torchia et al. 2016, Du et al. 2008).
   - Out of 534 methylation profiles of lower-grade glioma (LGG) (TCGA, Ceccarelli et al. 2016), preprocessed RNAseq data (FPKM) were available for 532 samples (Rahman et al. 2015).
   - Out of 155 methylation profiles of glioblastoma (GBM) (TCGA, Ceccarelli et al. 2016), preprocessed RNAseq data (FPKM) were available for 64 samples (Rahman et al. 2015).

2. Samples for which methylation profiles as well as immunohistological counts of TILs were available. 47 formalin-fixed paraffin-embedded (FFPE) tissue samples were sourced from the Institute of Neuropathology, University Medical Center Hamburg-Eppendorf (informed consent was obtained for all patients before analysis). For DNA extraction, samples were cut into 10 × 10 µm sections, and the ReliaPrep™ FFPE gDNA Miniprep System (Promega) was used. To make the methylation pattern detectable, approximately 100–500 ng of DNA was treated with bisulfite conversion by the EZ DNA Methylation Kit (Zymo Research). Subsequently, the DNA Clean & Concentrator-5 (Zymo Research) and the Infinium HD FFPE DNA Restore Kit (Illumina) were employed to clean and restore the DNA. Finally, the Infinium MethylationEPIC BeadChip Kit (Illumina) was utilized to quantify the methylation status of 850,000 CpG sites on an iScan device (Illumina). Preprocessing was similar to that applied to samples from public data repositories. Immunohistochemistry of FFPE tumor samples was conducted on an automated Ventana system, where anti-CD3 primary antibodies (Zytomed, M3974, dilution 1:100) were used for staining. CD3+ T cells were counted in three representative 2000 × 2000 pixel image regions at a magnification of ×400 for each sample.

All DNA methylation profiles were preprocessed using the R package minfi (Aryee et al. 2014), where single-sample normal-exponential out-of-band (noob) normalization was done on data, and Beta-scores were calculated for the final analysis (Ripley et al. 2020). CpG sites associated with single-nucleotide polymorphisms, sex chromosomes, and cross-reactive sites were excluded according to prior reports, which gave us a dataset of 428799 CpGs (Capper et al. 2018).

Clinical Samples

In addition to the training and validation samples, we used the validation set of the Heidelberg Brain Tumor Classifier (n = 1,104) to demonstrate the clinical applications of our algorithm (Capper et al. 2018). Clinical annotations were extracted from the supplementary material of the corresponding publication cited above.

1.2.2. Methods: Training

Two differential methylation analyses and a threshold analysis were conducted to obtain TIL signatures. This way, CpG sites specific to each lymphocyte type were identified and defined as the methylation signature of TILs. Finally, principal component analysis (PCA) was applied to these signatures, and the first principal component was regarded as the abundance score of the TILs. Figure 2 (Safaei et al. 2021) gives an overview of the above-mentioned procedures.

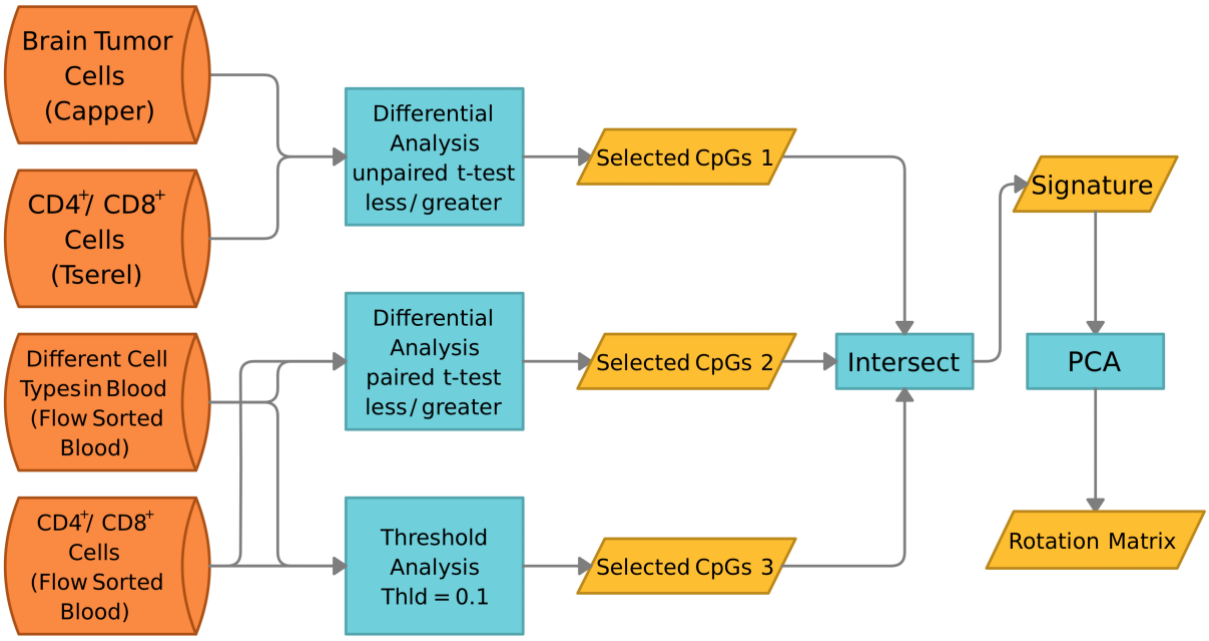


Figure 2. Overview of the method training (Safaei et al. 2021)

First, an unpaired one-sided Welch's t-test was employed to evaluate the significance of differential methylation between the profiles of 94 CD8+ T cells, 94 CD4+ T cells, as well as the combined set of both T cells (188 profiles) and the methylation profiles of brain tumors in the training cohort of Capper et al. (Capper et al. 2018, Jaffe 2020, Reinius et al. 2012). The significance of hypomethylation and hypermethylation was calculated for each of the 85 brain tumor categories individually. The average p-value for differential methylation over the 85 statistical tests was utilized to rank CpGs. Only CpGs with an average p-value < 0.05/428799 (Bonferroni correction) were retained for further analysis (selected CpGs 1 in Fig. 2). Secondly, a paired one-sided Welch's t-test was employed to evaluate the significance of differential methylation between the CD4+/CD8+ T cells and other peripheral blood cell types in a manner similar to the analysis of T cells and tumors (selected CpGs 2 in Fig. 2). Finally, those CpGs were selected, whose mean beta values were consistently higher or lower by a threshold of 0.1 in CD4+/CD8+ T cells compared to the remaining blood cell populations (selected CpGs 3 in Fig. 2). The final signatures are the intersection of selected CpGs 1, 2, and 3.

To use the obtained signatures in the TIL estimation of bulk tumor samples, principal component analysis was applied to them, which gave us a rotation matrix and a principal component. This principal component was then considered as the abundance score of the TILs. Figure 2 (Safaei et al. 2021) gives an overview of the conducted procedures.

1.2.3. Methods: Validation

To validate our method, two benchmarks were used: 1. immunohistological TIL counts and 2. gene expression-based TIL scores. To compare TIL scores, calculated from our method, with those from the benchmarks, the following two mathematical tools were utilized.

1. Spearman's rank correlation analysis was employed to determine the correlation coefficient (R) and p-value. P-values < 0.05 were considered statistically significant.

2. Robust linear regression, implemented in the R package MASS, was utilized to visualize the linear regression in scatter plot diagrams.

In our first benchmark, TIL counts were already measured during data acquisition. In our second benchmark, TIL scores were computed using well-established methods (Becht et al. 2016, Danaher et al. 2017). These methods were adapted for brain tumors by manually reviewing the signatures and excluding unspecific genes through correlation analyses (Bockmayr et al. 2019). TIL scores in our method were obtained by applying the learned PCA rotation matrix of each signature (DIME-CD4, DIME-CD8, DIME-TIL) on the validation data. Figure 3 (Safaei et al. 2021) provides an overview of the validation of our method with the second benchmark.

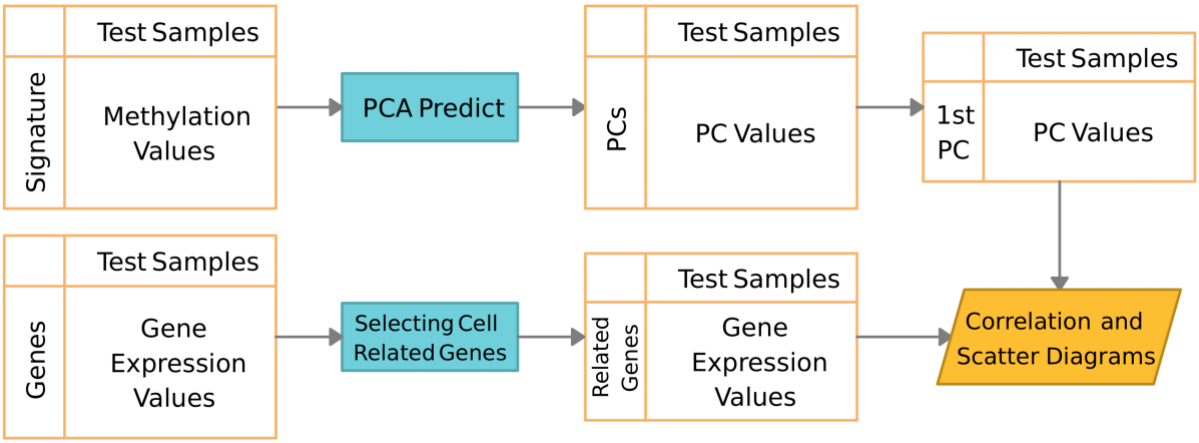


Figure 3. Overview of the method validation with the second benchmark (Safaei et al. 2021)

We have also studied how well our method correlates with the benchmark in comparison to two other previously published methods, namely MeTIL and methylCIBERSORT. To this end, we applied all three methods to the same datasets and performed correlation and regression analyses with the same benchmark. Finally, we compared correlation coefficients (R) and p-values to see which method has better performance. The MeTIL score was computed as described by Jeschke et al. (2017). MethylCIBERSORT analysis was performed as implemented by Chakravarthy et al. (2018) using the brain tumor-specific signature matrix published by Grabovska et al. (2020).

1.2.4. Methods: Statistical Analysis and Visualization of the Results

Gene ontology enrichment analyses were conducted using the *gometh* function (Phipson et al. 2016). Methylation-based diagnoses were calculated with the Heidelberg Brain Tumor Classifier (Capper et al. 2018). Differences in estimates of TILs between different tumor entities were computed with the Kruskal-Wallis test. Proportional hazard modeling was performed with the R package survival incorporating immune scores as continuous variables (Therneau and Grambsch 2000). P-values < 0.05 were considered statistically significant.

1.3. Results

The results of our study, in alignment with datasets and methodology, can also be classified into training outcomes and validation outcomes. Furthermore, the clinical function of our algorithm is demonstrated.

1.3.1. Training Results: TIL Signatures

As described in the section *Methods: Training*, the following three TIL signatures were obtained.

1. DIME-CD4 for CD4+ T cells,
2. DIME-CD8 for CD8+ T cells, and
3. DIME-TIL for CD4+ and CD8+ lymphocytes collectively.

These signatures consist of sets of CpGs, which exhibit either hypomethylation or hypermethylation in the corresponding T cell in contrast to tumor cells and other cell populations.

1. DIME-CD4 contains 4 hypermethylated and 26 hypomethylated CpGs,
2. DIME-CD8 contains 17 hypermethylated and 27 hypomethylated CpGs, and
3. DIME-TIL contains 105 hypermethylated and 92 hypomethylated CpGs.

Figure 4 (Safaei et al. 2021) visualizes these signatures via heatmaps. Each dot represents the methylation value of the corresponding CpG in the particular sample. Blue signifies hypomethylated values and red hypermethylated values. The methylation values partition the heatmaps into four distinct areas:

1. Red dots representing hypermethylated CpGs in the T cell, for which the signature is derived.
2. Blue dots representing hypomethylated CpGs in other cell lineages.
3. Blue dots representing hypomethylated CpGs in the T cell, for which the signature is derived.
4. Red dots representing hypermethylated CpGs in other cell lineages.

Figure 4d compares our signatures with MeTIL. In this comparison, methylation values of each signature hypermethylated CpGs were averaged for each cell type in the flow-sorted blood dataset. In contrast to our signature, CD19 B cells had the highest scores in MeTIL. Both methods could differentiate between lymphocytes and CD14 cells as well as granulocytes.

Table 2 (Safaei et al. 2021) presents pathways associated with lymphocyte-specific CpG signatures based on Gene Ontology enrichment analysis. The most significantly enriched pathways were:

1. For DIME-TIL, “regulation of innate immune response” (p = 3.51e-04, FDR = 0.64) and “T cell activation” (p = 5.86e-04, FDR = 0.64).
2. For DIME-CD4, “regulation of defense response to virus by virus” (p = 1.82e-05, FDR = 0.04) and “positive regulation of interleukin-2 biosynthetic process” (p = 1.19e-04, FDR = 0.13).
3. For DIME-CD8, “type I interferon signaling pathway” (p = 1.10e-04, FDR = 0.24) and “interferon-gamma-mediated signaling pathway” (p = 5.14e-04, FDR = 0.45).

Most of the enrichment outcomes did not remain significant after correction for multiple testing. Despite this, the most enriched categories are largely associated with the immune system. This shows that the identified CpGs are linked to immune system-related genes.

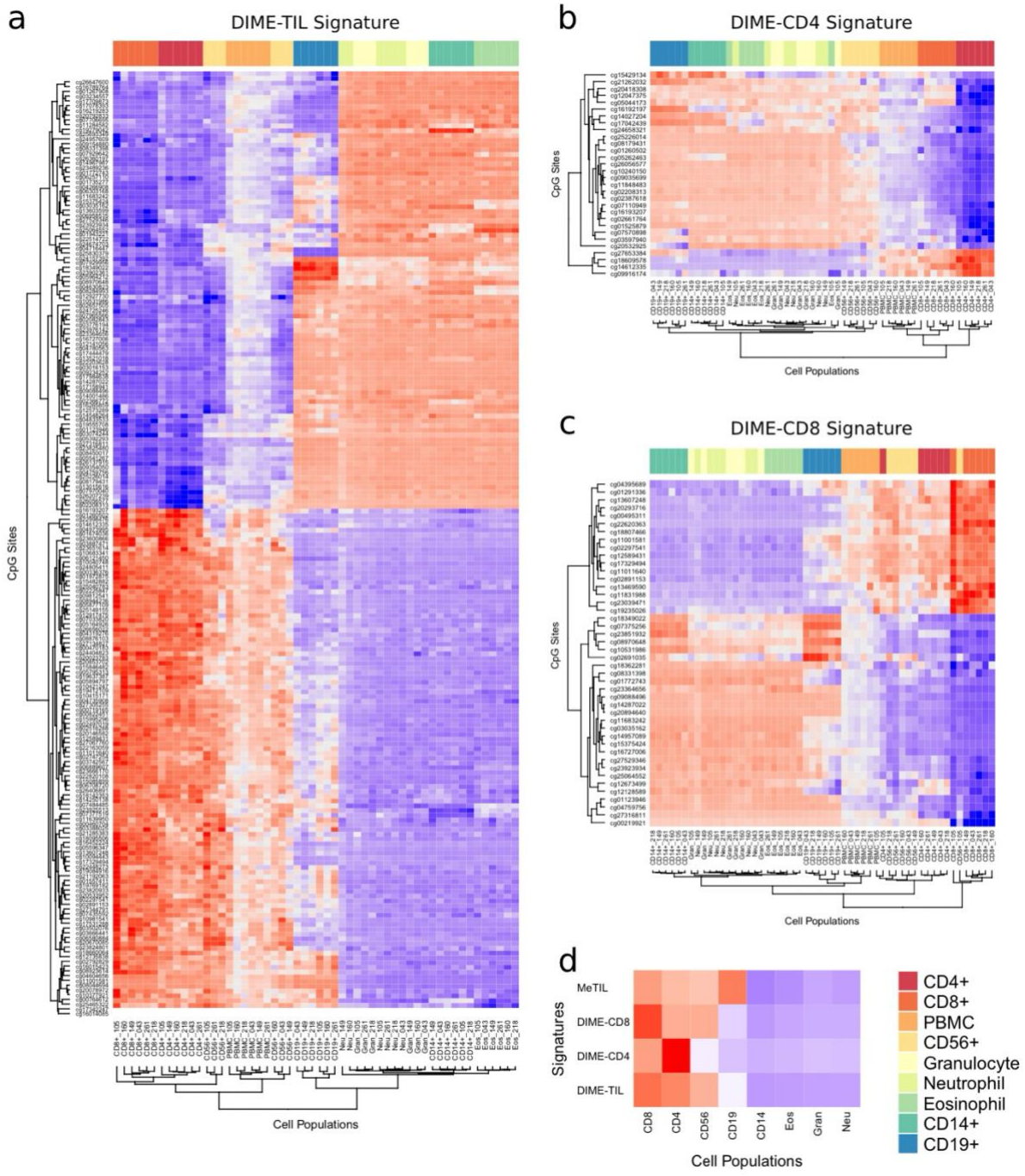


Figure 4. Visualization of the DIMEimmune signatures using heatmaps. a. DIME-TIL; b. DIME-CD4; c. DIME-CD8; d. Overview comparing calculated scores based on the three aforementioned signatures and MeTIL on the flow-sorted blood dataset. The DNA methylation values of samples of the same type have been averaged. (Safaei et al. 2021)

Cell Type	ID	TERM	DE	P.DE	FDR
TIL	GO:0045088	regulation of innate immune response	3	3.51E-04	0.64
	GO:0042110	T cell activation	4	5.86E-04	0.64
	GO:0060337	type I interferon signaling pathway	4	1.61E-03	1.00
	GO:0030217	T cell differentiation	3	4.00E-03	1.00
	GO:0031295	T cell costimulation	3	8.40E-03	1.00
	GO:1900017	positive regulation of cytokine production involved in inflammatory response	2	9.34E-03	1.00
	GO:0001816	cytokine production	2	9.54E-03	1.00
	GO:0030101	natural killer cell activation	2	9.86E-03	1.00
	GO:0043551	regulation of phosphatidylinositol 3-kinase activity	2	1.12E-02	1.00
	GO:0048535	lymph node development	2	1.16E-02	1.00
CD4	GO:0050690	regulation of defense response to virus by virus	3	1.82E-05	0.04
	GO:0045086	positive regulation of interleukin-2 biosynthetic process	2	1.19E-04	0.13
	GO:0006953	acute-phase response	2	4.76E-04	0.35
	GO:0050829	defense response to Gram-negative bacterium	2	1.20E-03	0.51
	GO:0033572	transferrin transport	2	1.27E-03	0.51
	GO:0034097	response to cytokine	2	1.56E-03	0.51
	GO:0042102	positive regulation of T cell proliferation	2	1.63E-03	0.51
	GO:0050731	positive regulation of peptidyl-tyrosine phosphorylation	2	7.97E-03	1.00
	GO:0033674	positive regulation of kinase activity	1	1.72E-02	1.00
	GO:0046039	GTP metabolic process	1	1.73E-02	1.00
CD8	GO:0060337	type I interferon signaling pathway	3	1.10E-04	0.24
	GO:0060333	interferon-gamma-mediated signaling pathway	3	5.14E-04	0.45
	GO:0043551	regulation of phosphatidylinositol 3-kinase activity	2	6.14E-04	0.45
	GO:0061512	protein localization to cilium	2	9.56E-04	0.52
	GO:0042110	T cell activation	2	3.09E-03	1.00
	GO:0048701	embryonic cranial skeleton morphogenesis	2	3.82E-03	1.00
	GO:0002479	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent	2	8.92E-03	1.00
	GO:0060071	Wnt signaling pathway, planar cell polarity pathway	2	1.59E-02	1.00
	GO:0019835	cytolysis	1	1.81E-02	1.00
	GO:0032897	negative regulation of viral transcription	1	1.97E-02	1.00

Table 2. Enrichment analysis of immune signatures (Safaei et al. 2021)

1.3.2. Training Results: PCA Rotation Matrix

To estimate TIL in bulk samples, besides the TIL signatures, the PCA rotation matrix is also required. TIL scores of a new bulk sample can then be calculated by applying the PCA rotation matrix to the methylation values of the CpGs indicated by the TIL signature under study. The orientation of the first principal component was chosen to correlate with the negative mean of hypomethylated sites. The rotation matrix is a pure mathematical tool and has no biological meaning. Therefore, further visualization is not needed.

1.3.3. Validation Results: Immunohistochemical Benchmark

Figure 5 (Safaei et al. 2021) presents the validation results, which compare our TIL scores with the ones from our first benchmark (see section *Methods: Validation*). Figure 5a-c depicts examples of cases with low/absent, intermediate, and high infiltration of CD3+ cells. The strongest correlation with the number of CD3+ cells was observed with the DIME-TIL score (R = 0.74, p = 2e-09, Figure 5d), followed by MeTIL (R = 0.32, p = 0.029, Figure 5e). A positive correlation was also observed with the MethylCIBERSORT TIL scores (R = 0.26), however it did not reach statistical significance.

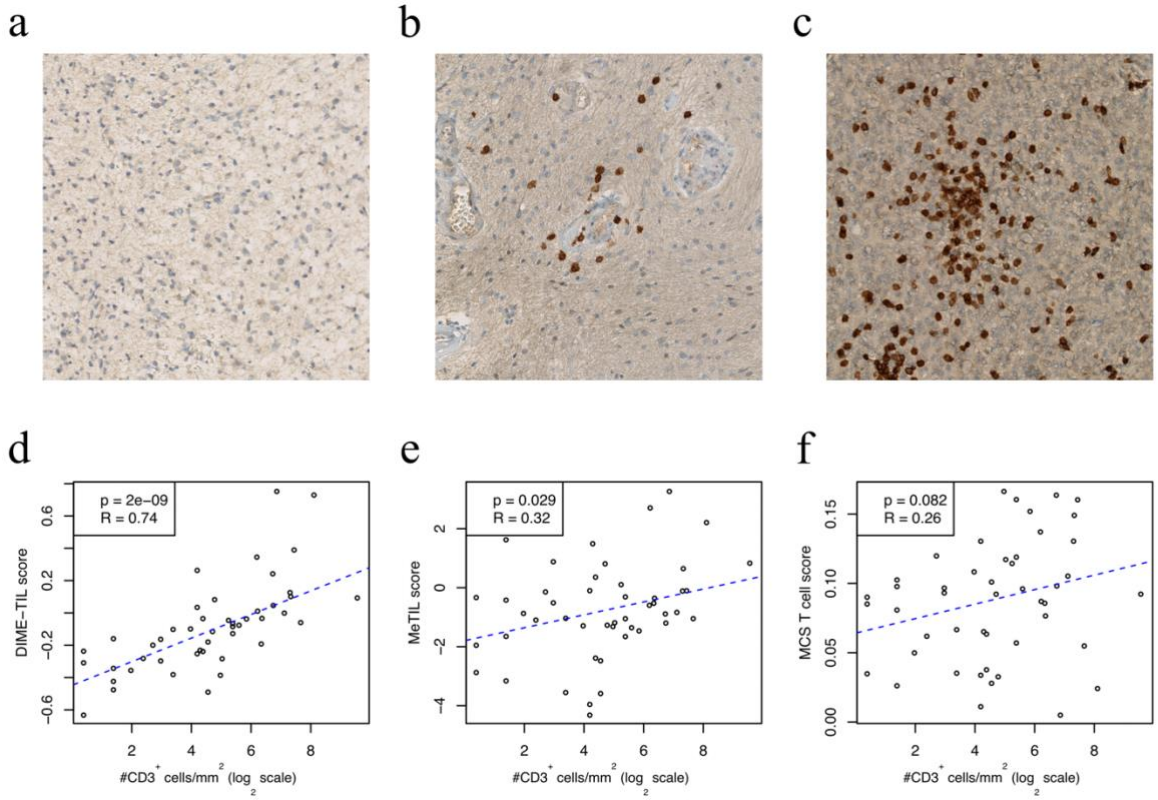


Figure 5. Immunohistological validation of methylation-based TIL estimates in a cohort of 47 brain tumors. Examples of tumors with low/absent CD3+ cells: a. anaplastic astrocytoma IDH-mutant, WHO grade III, intermediate CD3+ cells; b. rosette-forming glioneuronal tumor, WHO grade I, and high numbers of CD3+ cells; c. glioblastoma IDH wild-type, WHO grade IV. Immunohistological counts of CD3+ cells are compared to DIME-TIL (d), MeTIL (e), and the sum of the T cell scores analyzed by MethylCIBERSORT (f). (Safaei et al. 2021)

1.3.4. Validation Results: Gene Expression Benchmark

Figures 6 to 8 (Safaei et al. 2021) present the validation results, comparing our TIL scores with the ones from our second benchmark (see section *Methods: Validation*).

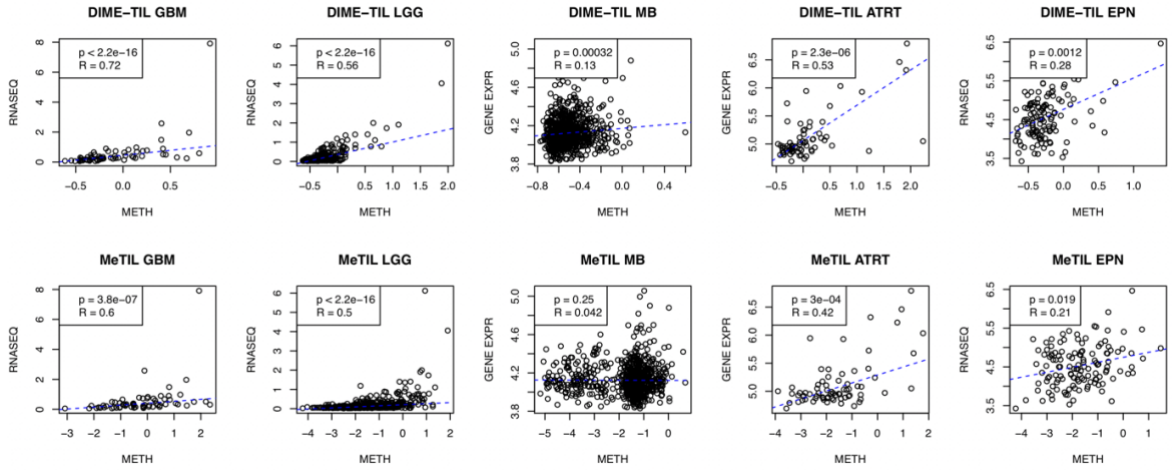


Figure 6. Comparison of DIME-TIL and MeTIL with gene expression/RNAseq-based estimation of T cells (Safaei et al. 2021)

Figure 6 (Safaei et al. 2021) depicts the comparison between the correlations of DIMEimmune and MeTIL with the gene expression-based method. The correlations are ranked identically for both methods as presented in Table 3.

	R (DIME-TIL)	R (METIL)
GBM	0.72	0.60
LGG	0.56	0.50
ATRT	0.53	0.42
EPN	0.28	0.21
MB	0.13	N.S.

Table 3. Correlations of DIMEimmune and MeTIL with gene expression-based method. N.S. = not significant.

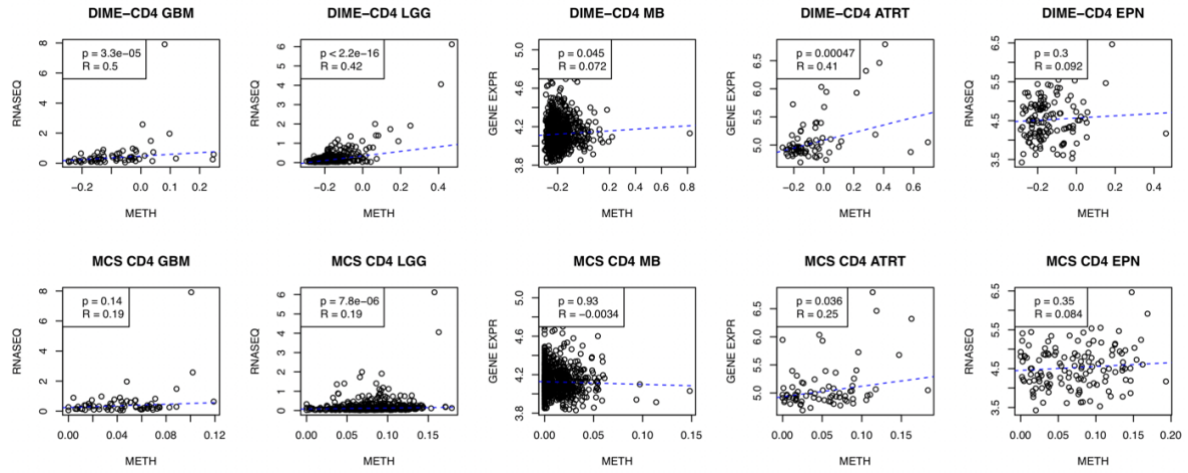


Figure 7. Comparison of DIME-CD4 and CD4+ T cell estimates obtained from MethylCIBERSORT (MCS) with gene expression/RNAseq-based estimation of CD4+ T cells (Safaei et al. 2021)

Figure 7 (Safaei et al. 2021) presents the comparison between the correlations of DIME-CD4 and CD4 estimations derived from MethylCIBERSORT ("CD4_Eff" + "Treg") with the gene expression-based estimation of CD4+ T cells. The correlations are presented in Table 4.

	R (DIME-CD4)	R (MCS CD4)
GBM	0.50	N.S.
LGG	0.42	0.19
ATRT	0.41	0.25
MB	0.07	N.S.
EPN	N.S.	N.S.

Table 4. Correlations of DIME-CD4 and MethylCIBERSORT with gene expression-based method. N.S. = not significant.

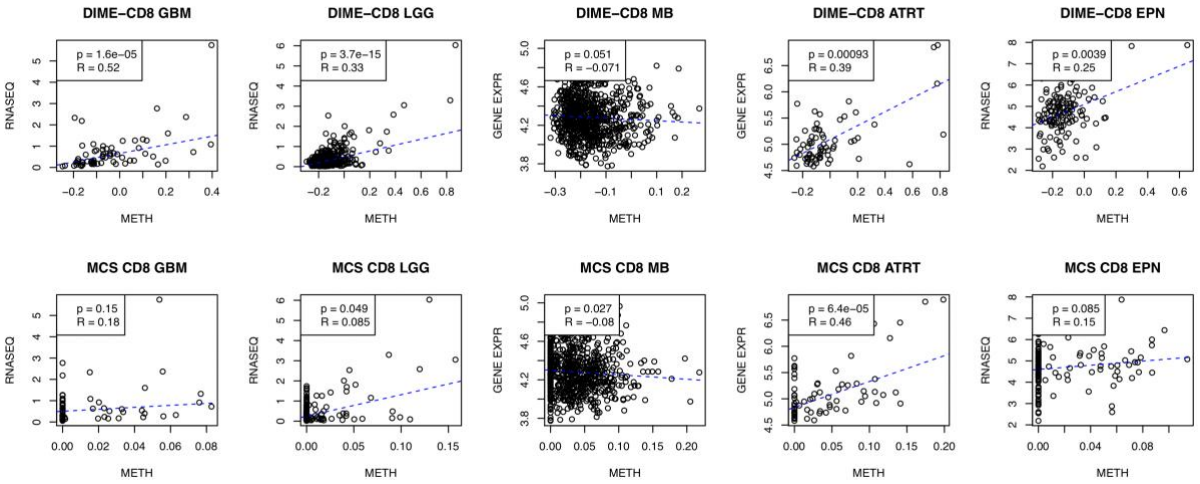


Figure 8. Comparison of DIME-CD8 and CD8+ T cell estimates obtained from MethylCIBERSORT with gene expression-based estimation of CD8+ T cells. (Safaei et al. 2021)

Figure 8 (Safaei et al. 2021) depicts the comparison between the correlations of DIME-CD8 and CD8+ estimations derived from MethylCIBERSORT with the gene expression-based estimation of CD8+ T cells. The correlations are presented in Table 5.

	R (DIME-CD8)	R (MCS CD8)
GBM	0.52	N.S.
ATRT	0.39	0.46
LGG	0.33	0.09
EPN	0.25	N.S.
MB	N.S.	-0.08

Table 5. Correlations of DIME-CD8 and MethylCIBERSORT with gene expression-based method. N.S. = not significant.

Overall, the results of our method correlate better with those of well-established gene expression-based results in all brain tumor types compared to previously published algorithms, except for the estimation of CD8+ T cells in ATRT.

1.3.5. Clinical Relevance and Association with Survival

Figure 9 (Safaei et al. 2021) presents the TIL scores derived from applying our method to the validation set of the Heidelberg Brain Tumor Classifier (Capper et al. 2018). As anticipated, the highest TIL score was detected in CNS lymphoma (LYMPHO). High TIL scores were also found in inflammatory and reactive control tissue (CONTR, REACT and CONTR, INFLAM). Other entities with high TIL scores included mesenchymal glioblastoma (GBM, MES), MYC ATRT, anaplastic pilocytic astrocytoma (ANA PA), melanoma (MELAN), chordoma (CHORDM), and pituitary adenomas (PITAD, TSH). Conversely, low TIL scores were found in posterior fossa B ependymoma (EPN, PF B), paraganglioma (PGG, nC), and group 4 medulloblastoma (MB, G4).

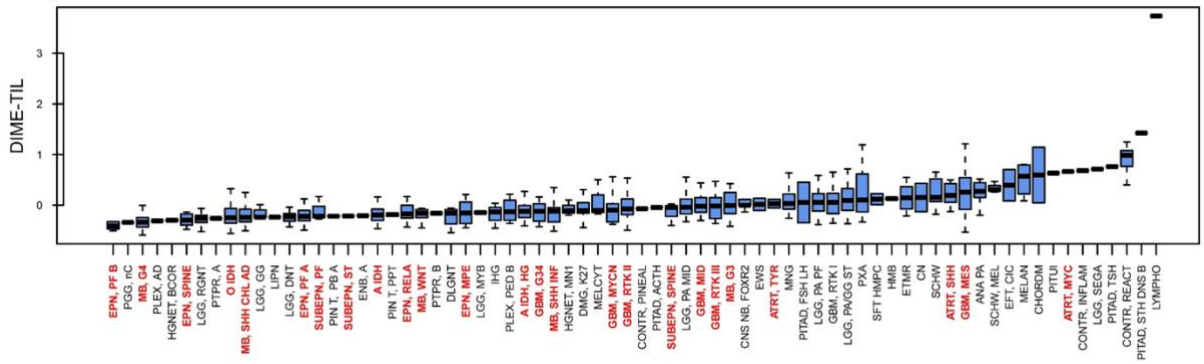


Figure 9. The estimated score of TIL for different subgroups of the Capper et al. validation data set. (Safaei et al. 2021)

Figure 10 (Safaei et al. 2021) presents the distribution of the via our method calculated TIL scores across brain tumor subgroups (GBM/LGG, MB, ATRT, and EPN) of the validation set of the Heidelberg Brain Tumor Classifier (Capper et al. 2018). Since WHO grade III and IV tumors were present in both the LGG and the GBM datasets, these datasets were combined into a single glioma dataset. Methylation-based subgroup diagnoses were determined using the Heidelberg Brain Tumor Classifier and the p-values were calculated based on the Kruskal-Wallis test.

a. **Glioma**: Among the glioma subgroups, the TIL scores were significantly different (p = 1.6e-36, Figure 10a). Estimated infiltration was lowest in IDH (Isocitrate dehydrogenase) mutated oligodendroglioma and IDH mutated astrocytoma, and highest in mesenchymal glioblastoma.
b. **Medulloblastoma**: Although statistically significant differences were observed between medulloblastoma subgroups (p = 1.7e-49, Figure 10b), the overall TIL score was low. Estimated infiltration was lowest in Group 4 medulloblastoma, and highest in Group 3.
c. **ATRT**: ATRTs showed noticeable lymphocytic infiltration. The ATRT MYC subgroup had significantly more TIL scores than the TYR (intermediate) and SHH (lowest) subgroups (p = 0.00012, Figure 10c).
d. **Ependymoma**: Among the ependymoma subgroups, the TIL scores were significantly different (p = 4.4e-0.8, Figure 10d). Estimated infiltration was lowest in the YAP and the PF B subgroups, and highest in the PF A and the RELA subgroups.

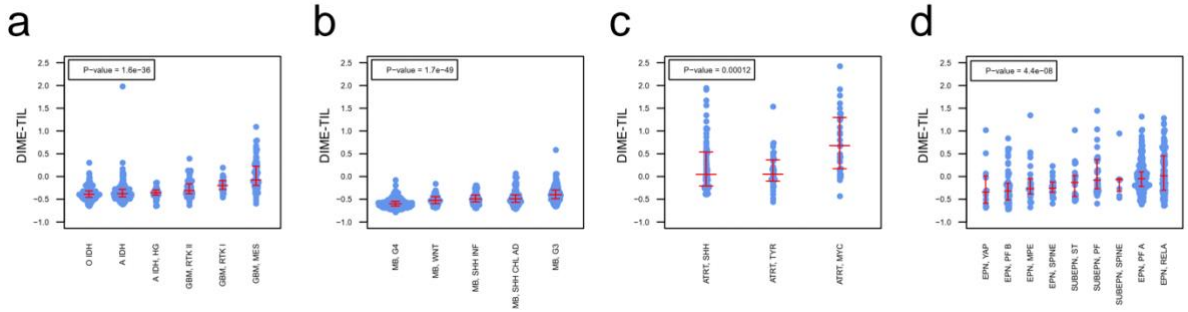


Figure 10. Estimated TIL scores for glioma (a), medulloblastoma (b), ATRT (c), and ependymoma (d) samples from our validation data set (Table 1). (Safaei et al. 2021)

Figure 11 (Safaei et al. 2021) illustrates associations between TIL scores and overall survival in glioma, medulloblastoma, ATRT, and ependymoma.

a. **Glioma**: The DIME-TIL score was identified as a strong negative prognostic factor (p = 6.06e-15, HR = 6.26, Figure 11a). However, within methylation-based subgroups, no significant association with survival was observed.
b. **Medulloblastoma**: The DIME-TIL score was identified as a negative prognostic factor (p = 0.003, HR = 4.3, Figure 11b). However, within methylation-based subgroups, no significant association with survival was observed.
c. **ATRT**: The DIME-TIL scores were not associated with survival in ATRT and its studied subgroups, possibly due to the smaller sample size compared to other tumor types (Figure 11c).
d. **Ependymoma**: The DIME-TIL scores were not associated with survival in ependymoma and its studied subgroups, despite sufficient survival data (Figure 11d).

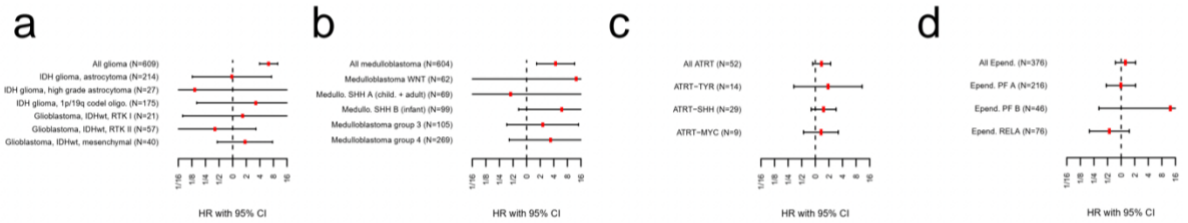


Figure 11. Survival analysis for prognostic relevance of DIME-TIL scores based on the same dataset as Figure 10, presented as forest plots. HR: hazard ratio, CI: confidence interval. (Safaei et al. 2021)

1.4. Discussion

Although TIL estimation methods based on immunohistology (Bienkowski and Preusser 2015) and gene-expression data (Bockmayr et al. 2018, Bockmayr et al. 2019, Griesinger et al. 2015, Gentles et al. 2015, Wang et al. 2017) are well established, approaches using methylation data for TIL estimation remain uncommon. On the other hand, several institutions use methylation data in the brain tumor diagnostic routine (Capper et al. 2018, Pickles et al. 2020). To address this gap, we introduced a robust method for TIL estimation based on bulk methylation data.

1.4.1. Methylation-Based Methods and Gene Expression

In medulloblastoma, all algorithms (MethylCIBERSORT, MeTIL, and DIMEimmune) showed weak correlations with gene expression-based approaches, especially for CD8+ T cells. Among these, only DIME-TIL ($R = 0.13$, $p = 0.00032$) and DIME-CD4 ($R = 0.072$, $p = 0.045$) showed significant correlations. The low number of tumor-infiltrating lymphocytes in medulloblastoma, as shown in Figure 9 (Safaei et al. 2021), may lead to a poor signal-to-noise ratio. Therefore, immunohistology, gene expression (particularly RNA-sequencing), or single-cell sequencing methods may be more suitable for analyzing tumors with very low amounts of infiltrating lymphocytes. Nevertheless, methylation-based analysis remains advantageous due to its broader applicability.

1.4.2. DIMEimmune and MeTIL

The MeTIL algorithm was developed through differential methylation analysis between lymphocytes and breast cancer tumor cells (Jeschke et al. 2017). While it has demonstrated robust results for breast cancer, our approach had better performance in brain tumors. Specifically, the TIL scores obtained from MeTIL for medulloblastoma in Figure 6 (Safaei et al. 2021) show a bimodal distribution. This could be due to a nonspecific signature in medulloblastoma and emphasizes the necessity of a brain tumor-specific algorithm such as DIMEimmune.

1.4.3. DIMEimmune and MethylCIBERSORT

In the correlation analysis with gene expression data, DIME-CD8 and DIME-CD4 performed better than MethylCIBERSORT except in estimating CD8+ and CD4+ T cells in ATRT. This can be explained by the fact that the tumor reference used by Grabovska et al. for MethylCIBERSORT included only cell lines from medulloblastoma and rhabdoid tumors (Grabovska et al. 2020). This shows the importance of using tumor-specific references and suggests that our approach is expected to estimate TIL abundance in CNS tumors more precisely, particularly in rare tumor types, for which only few or no cell line data are available.

1.4.4. DIME-TIL Scores and Previously Reported Findings

Overall, estimates of TIL abundance based on DIME-TIL align with previously reported findings.

- High TIL scores in CNS lymphoma and reactive and inflammatory tissues support the validity of this method.
- Lower TIL scores in low-grade gliomas and IDH-mutated gliomas compared to IDH wild-type gliomas, are consistent with earlier results (Berghoff et al. 2017, Lohr et al. 2011, Weenink et al. 2019). Additionally, the mesenchymal subgroup is well-known for high TIL scores (Bockmayr et al. 2019, Rutledge et al. 2013).
- Medulloblastoma showed relatively low TIL scores, which is consistent with previous studies (Bockmayr et al. 2018, Vermeulen et al. 2017). However, the methylation-based analysis did not find larger amounts of T cells in the SHH subgroups as previously identified by gene expression analysis (Bockmayr et al. 2018, Margol et al. 2015). This could be due to a high noise-to-signal ratio.

- Higher TIL estimates in MYC ATRTs compared to SHH and TYR ATRTs have also been previously reported (Leruste et al. 2019, Chun et al. 2019).
- Among the ependymoma subgroups, PF A ependymoma had the highest estimated number of TILs, which is compatible with the inflammatory phenotype characteristic of this subtype (Griesinger et al. 2015, Hoffman et al. 2014).

1.4.5. Survival Analysis

Survival analysis demonstrated a significant negative prognostic impact of the DIME-TIL score in both the overall glioma and medulloblastoma cohorts, although this effect was not observed within methylation subgroups.

- In glioma, this negative prognostic effect may be due to the higher number of TILs in more aggressive WHO grade IV gliomas (Lohr et al. 2011).
- Similarly, the variance in TIL estimates among medulloblastoma subgroups, particularly the higher levels in Group 3 medulloblastoma, could act as a confounding factor.

The literature presents conflicting reports on the prognostic significance of tumor-infiltrating lymphocytes in glioma and medulloblastoma (Bockmayr et al. 2018, Bockmayr et al. 2019, Grabovska et al. 2020, Han et al. 2014, Lohr et al. 2011, Marinari et al. 2020, Rutledge et al. 2013, Vermeulen et al. 2017). These discrepancies show the need for more research on the prognostic role of TILs in brain tumors, using large cohorts with well-defined diagnostic subgroups. Since the histological or molecular subtype might affect the results, methylation data-based methods offer a valuable option for such studies due to their widespread availability.

1.5. Reception and Applications

DIMEimmune has been employed in various research projects. For instance, it has been used in the CD4+ and CD8+ T cells estimation to find out how lymphocyte dynamics may change between primary ATRT tumors and recurrences (Johann et al. 2023). In another study, DIMEimmune was used in CD4+ and CD8+ T cell estimation in peripheral nerve sheath tumors and showed that atypical neurofibromas have the highest lymphocyte infiltration (Kresbach et al. 2023). Furthermore, DIMEimmune revealed significant immune infiltration differences between MPE-A and MPE-B subtypes of myxopapillary ependymomas (Bockmayr et al. 2022). As of the completion of this doctoral thesis, our paper has been cited in 21 publications (Google Scholar, 21.02.2025).

1.6. Conclusion

Methylation data is increasingly used in tumor diagnostics, but there are still limited methods to extract information about the tumor microenvironment from this data type. To address this gap, we developed DIMEimmune (Differential Methylation Analysis for Immune Cell Estimation) using differential methylation and principal component analysis. It estimates the abundance of CD4+ and CD8+ T cells from bulk methylation data. When compared to other methods such as MethylCIBERSORT and MeTIL, DIMEimmune showed stronger correlations with both gene expression-based and immunohistological results across different brain tumor

types. This method has the potential to be used as an important prognostic or predictive tool in future CNS tumor research.

2. Original Paper

ONCOIMMUNOLOGY
2021, VOL. 10, NO. 1, e1932365 (14 pages)
https://doi.org/10.1080/2162402X.2021.1932365

Taylor & Francis
Taylor & Francis Group

ORIGINAL RESEARCH

OPEN ACCESS Check for updates

DIMEimmune: Robust estimation of infiltrating lymphocytes in CNS tumors from DNA methylation profiles

Sepehr Safaei[a,b,c], Malte Mohme[d], Judith Niesen[a,b,e], Ulrich Schüller[a,b,c]*, and Michael Bockmayr[a,b,e,f]*

[a]Department of Pediatric Hematology and Oncology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [b]Research Institute Children's Cancer Center Hamburg, Hamburg, Germany; [c]Institute of Neuropathology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [d]Department of Neurosurgery, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [e]Mildred Scheel Cancer Career Center HaTriCS[4], University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [f]Institute of Pathology, Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin and Berlin Institute of Health, Berlin, Germany

ABSTRACT

The interaction of CNS tumors with infiltrating lymphocytes plays an important role in their initiation and progression and might be related to therapeutic responses. Gene expression-based methods have been successfully used to characterize the tumor microenvironment. However, methylation data are now increasingly used for molecular diagnostics and there are currently only few methods to infer information about the microenvironment from this data type. Using an approach based on differential methylation and principal component analysis, we developed DIMEimmune (Differential Methylation Analysis for Immune Cell Estimation) to estimate $CD4^+$ and $CD8^+$ T cell abundance as well as tumor-infiltrating lymphocytes (TILs) scores from bulk methylation data. Well-established approaches based on gene expression data and immunohistochemistry-based lymphocyte counts were used as benchmarks. The comparison of DIMEimmune to the previously published MethylCIBERSORT and MeTIL algorithms showed an improved correlation with both gene expression-based and immunohistological results across different brain tumor types. Further, we applied our method to large datasets of glioma, medulloblastoma, atypical teratoid/rhabdoid tumors (ATRTs) and ependymoma. High-grade gliomas showed higher scores of tumor-infiltrating lymphocytes than lower-grade gliomas. There were overall only few tumor-infiltrating lymphocytes in medulloblastoma subgroups. ATRTs were highly infiltrated by lymphocytes, most prominently in the MYC subgroup. DIMEimmune-based estimates of TILs were a significant prognostic factor in the overall cohort of gliomas and medulloblastomas, but not within methylation-based diagnostic subgroups. To conclude, DIMEimmune allows for robust estimates of TIL abundance and might contribute to establishing them as a prognostic or predictive factor in future studies of CNS tumors.

ARTICLE HISTORY
Received 21 January 2021
Revised 23 April 2021
Accepted 16 May 2021

KEYWORDS
Brain tumor; microenvironment; DNA methylation; immune cells; tumor-infiltrating lymphocytes

Introduction

The immune microenvironment is a key factor for tumor growth and progression in various cancers including central nervous system (CNS) tumors.[1,2] There has been an increasing interest in tumor immunology with the avenue of immunotherapies, as there is growing evidence that the tumor microenvironment influences the therapeutical outcome. Tumor-infiltrating lymphocytes (TILs) have been established as an important predictive and prognostic biomarker in several solid tumors, particularly in breast cancer (see[3–5] for review). In neuro-oncology, TILs have been primarily studied in glioblastoma and were shown to be associated with molecular alterations, such as *NF1* and *RB1* mutations, although the reported prognostic associations are still being controversially discussed.[6–8] Using transcriptomic approaches, immunological differences between molecular subgroups were identified in medulloblastoma (MB), ependymoma and ATRT.[9–11] Recently, profiling the immune microenvironment of over 6,000 primarily pediatric brain tumors (medulloblastomas, malignant rhabdoid tumors, and high-grade gliomas) using MethylCIBERSORT, Grabovska *et al*. showed associations of particular immune cells with molecular subgroups, mutations, as well as the overall survival in these entities.[12,13]

Different strategies have emerged for the quantification of tumor-infiltrating immune cells. TILs are frequently scored on hematoxylin and eosin (HE) stained slides, which is easily possible in e.g. breast cancer, lung cancer or glioblastoma.[3,4] Recently, machine learning methods have been applied to optimize morphological TIL quantification.[14] However, it becomes increasingly difficult in tumors with morphological similarities to immune cells, like small-round-blue-cells tumors, where immunohistochemical analyses are needed. Also, the high diversity of rare brain tumor entities, spatial heterogeneity, and potential interrater variations make it difficult to obtain solid TIL phenotyping from tissue sections. Utilizing molecular diagnostics such as methylation data to determine the degree of infiltration and subset phenotype of immune cells is therefore appreciated.

CONTACT Ulrich Schüller u.schueller@uke.de; Michael Bockmayr m.bockmayr@uke.de Research Institute Children's Cancer Center, Hamburg, Martinistr. 52, N63, 20251 Hamburg, Germany
*These authors contributed equally
Supplemental data for this article can be accessed on the publisher's website

Various techniques have been developed to estimate tumor-infiltrating immune cells from bulk molecular profiling data. Deconvolution algorithms based on transcriptomic techniques are most widely used and include approaches based on specific gene signatures for microenvironment cell populations and the CIBERSORT method.[15–17] On the one hand, estimates based on specific immune cell signatures are usually defined as the average expression value of the cell-type-specific genes. However, most published signatures were not optimized for brain tumors resulting in potentially unspecific results. Therefore, we recently optimized these methods allowing for robust estimates of microenvironmental cells in medulloblastoma and pediatric high-grade glioma.[8,9]

On the other hand, CIBERSORT uses a support-vector regression-based deconvolution algorithm to quantify the relative amount of 22 immune subpopulations, which could be used to identify prognostic markers in various cancers.[18] The algorithm provides a *p*-value for each sample, which is a measure of confidence in the results. In a large gene expression-based meta-analysis of breast cancer, the overall amount of immune infiltration was negatively correlated with this *p*-value, indicating that lower amounts of tumor-infiltrating immune cells might result in less robust deconvolution results.[19]

These transcriptomic approaches are well established and validated and led to several new insights on the immunological landscape of solid tumors. However, transcriptomic data are only rarely used in routine diagnostics of brain tumors. Conversely, global DNA methylation analysis has been extensively used for molecular analyses of brain tumors,[20–24] and there is an increasing number of applications in the clinical routine for the diagnostic classification of brain tumors and other neoplasms.[25–28] Therefore, methods that can robustly quantify immune cells based on DNA methylation are needed and currently much more widely applicable in neuro-oncology than transcriptomic approaches.

Recently, DNA methylation-based algorithms have been established to study tumor-infiltrating immune cells. Jeschke *et al.* used differential methylation analysis between lymphocytes and breast cancer to establish a methylation signature (MeTIL) that recapitulates TIL evaluations and their prognostic value in breast cancer.[29] Adapting the CIBERSORT algorithm to DNA methylation data, Chakravarthy *et al.* introduced MethylCIBERSORT. Using methylation profiles from head and neck squamous cell carcinoma, they divided their data into immune hot tumors with better response and immune cold tumors with worse response to the therapy.[13] As previously described, this approach has been recently applied by Grabovska *et al.* to study a large cohort of brain tumors.[12] A prerequisite for the reliable immune cell estimation with MethylCIBERSORT is a tumor reference signature for the entity under consideration. Therefore, cell line data have been used, which is, however, not available for rare brain tumor entities. Grabovska *et al.* therefore used 25 cell lines from rhabdoid tumors and medulloblastoma only as reference.[12]

As nonspecific tumor signatures result in nonreliable deconvolution results, and as the reference data are unavailable for many rare brain tumor entities, we established DIMEimmune (Differential Methylation Analysis for Immune Cell Estimation), a robust DNA methylation-based method for the quantification of TILs (DIME-TIL), $CD4^+$ (DIME-CD4) and $CD8^+$ (DIME-CD8) lymphocytes in central nervous system tumors. The method is not relying on a reference signature of pure tumor cells from the studied tumor entity and can therefore be directly applied to any CNS tumor.

Materials and Methods

All data analyses were performed using the statistical programming language R version 3.6.0[30] with the packages minfi, lumi, missMethyl, ComplexHeatmap, MASS, beeswarm, TCGAbiolinks, MethylCIBERSORT, FlowSorted.Blood.450k and survival.[13,31–39]

Datasets and preprocessing of previously published methylation data

The methylation data used to train our method as well as the data used for comparison with gene expression-based deconvolution algorithms were previously published, analyzed on the Illumina Infinium Methylation 450K Bead Chip and are available from public data repositories. These raw data can be divided into training and validation data (Figure 1a). As reference for immune cells, publicly available methylation profiles of magnetic-activated cell sorted $CD4^+$ and $CD8^+$ lymphocytes from 101 cases were downloaded from Gene Expression Omnibus (GEO).[40,41] Not all samples contain data for both $CD4^+$ and $CD8^+$ lymphocytes, therefore profiles from patients with missing data either for $CD4^+$ or $CD8^+$ lymphocytes were excluded resulting in a dataset of 94 samples. The training cohort for the brain tumor classifier published by Capper *et al.* was used as a reference for CNS tumors.[25] As we computed differential methylation between bulk tumor methylation profiles and immune cells, samples from reference tissue and tumors with expected high immune infiltration were excluded (methylation classes: LYMPHO, PLASMA, CONTR INFLAM, CONTR REACT, MELAN, MELCYT) resulting in a dataset of 2706 samples from 85 diagnostic categories. As validation datasets, the validation cohort from Capper *et al.* was used (1104 samples), as well as 763 methylation profiles from medulloblastoma (Cavalli *et al.*[20]), 162 methylation profiles from atypical teratoid/rhabdoid tumor (ATRT),[23,42] 534 methylation profiles from lower-grade glioma (LGG) (TCGA[43]), 155 samples from glioblastoma (GBM) (TCGA[43]), and 557 samples from ependymoma (EPN) (Pajtler *et al.*).[22] Clinical annotations were extracted from the supplementary material of the corresponding publications cited above. All DNA methylation profiles were preprocessed with the R-package minfi[31] using single-sample normal-exponential out-of-band (noob) normalization, beta-scores were used for the final analysis.[38] CpG sites associated with single-nucleotide polymorphisms, sex chromosomes, and cross-reactive sites were excluded as previously reported, resulting in a dataset of 428799 CpGs.[25]

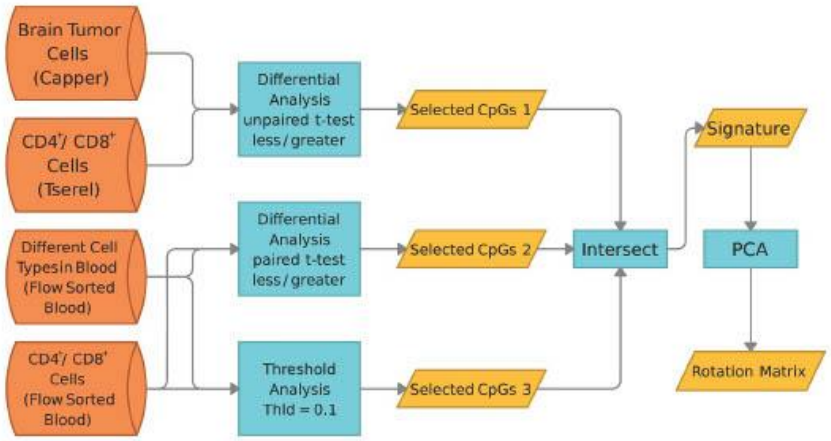
Matching gene expression data were obtained for medulloblastoma, ATRT, glioma and ependymoma from GEO or the

a

Dataset Type	Cell Type	ID	Author	Number of Cases
Training Dataset	Brain Tumor (85 Entities)	GSE90496	Capper 2018	2706
	$CD4^+$	GSE59065	Tserel 2015	94
	$CD8^+$	GSE59065	Tserel 2015	94
	Different Cells in Blood	GSE35069	Reinius 2012	60
Validation Dataset	Medulloblastoma	GSE85212	Cavalli 2018	763 (763)*
	Ependymoma	GSE64415	Pajtler 2015	557 (129)*
	ATRT	27960086	Torchia 2017	162 (88)*
	Low Grade Glioma	TCGA-LGG		534 (532)*
	Glioblastoma	TCGA-GBM		155 (64)*

* Number of samples for which both RNASeq / gene expression and methylation data are available.

b



c

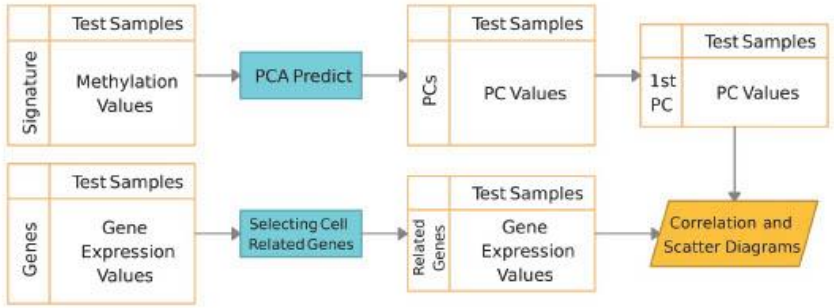


Figure 1. Overview of the data and the DIMEImmune method. a. Data sources and number of samples; b. Overview of the method: the final lymphocyte-specific signature is the intersection of three sets of CpGs. The first two sets are obtained within the framework of differential methylation analysis and the third one by threshold analysis. The first group is selected based on the difference between brain tumor samples, and $CD4^+$ and $CD8^+$ T cells by applying an unpaired *t*-test. The second group is selected by applying a paired *t*-test on different immune cell types in the blood. Additionally, a threshold analysis is conducted, which gives the third group of CpGs. The final estimate for lymphocyte infiltration is obtained from dimensionality reduction using PCA; c. Validation: Methylation-based estimates are obtained by applying the learned rotation matrix of PCA on the validation data. Gene expression-based estimates are computed from specific signatures as previously described.[8] Both methods are compared on the same samples using RNAseq/gene expression and methylation-based values in the form of correlation and scatter diagrams.

European Genome-phenome Archive. Expression data for medulloblastoma (763 samples) were analyzed on the Affymetrix Human Gene 1.1 ST Array and preprocessed as previously described.[9] Expression data for ATRT (88 samples)

were analyzed on the Illumina HT12 gene expression array and preprocessed with the R-package lumi.[23,39] Preprocessed RNAseq data (FPKM) for LGG (532 samples) and GBM (64 samples) were obtained from Rhaman *et al.*[44] Expression data for EPN (129 samples) were analyzed on the Affymetrix HG U133 Plus 2.0 microarray and preprocessed with the R-package affy using the custom chip definition file hgu133plus2hsentrezgcdf (v19.0.0).[45,46]

Immunohistological validation

Parallel methylation and immunohistological analysis were performed on 47 diagnostic cases including pituitary adenomas, gliomas, MBs, EPNs and ATRTs. Formalin-fixed paraffin embedded (FFPE) tissue was obtained from the archives of the Institute of Neuropathology, University Medical Center Hamburg-Eppendorf. Informed consent was obtained for all patients prior to the analysis.

For methylation analysis, DNA isolation was performed on FFPE tissue, 10 × 10 µm sections were cut and DNA isolated using the ReliaPrep™ FFPE gDNA Miniprep System (Promega) according to manufacturer's instructions. About 100–500 ng DNA was used for bisulfite conversion by the EZ DNA Methylation Kit (Zymo Research). Afterward, the DNA Clean & Concentrator-5 (Zymo Research) and the Infinium HD FFPE DNA Restore Kit (Illumina) were employed to clean and restore the converted DNA. Finally, the Infinium MethylationEPIC BeadChip Kit (Illumina) was used to quantify the methylation status of 850,000 CpG sites on an iScan device (Illumina). Preprocessing was performed analogously to samples from public data repositories. FFPE human tumor samples were used for immunohistochemistry. Analyses were done on an automated Ventana system using anti-CD3 primary antibodies (Zytomed, M3974, 1:100). $CD3^+$ cells were counted in three representative image regions of 2000 × 2000 pixels (at magnification x400) for each sample.

DNA methylation-based estimation of tumor-infiltrating lymphocytes

Three signatures were defined to estimate the amount of tumor-infiltrating $CD8^+$ T cells (DIME-CD8), tumor-infiltrating $CD4^+$ T cells (DIME-CD4), and a mixed signature of tumor-infiltrating lymphocytes (DIME-TIL). To this end, multiple steps of differential methylation analysis were applied to identify specific CpG sites. This was followed by dimensionality reduction using principal component analysis (PCA) resulting in an estimate for the studied cell populations. The details of the method are described in the following.

Differential methylation analysis between immune cells and CNS tumors

First, an unpaired one-sided Welch's *t*-test was used to assess the significance of differential methylation between the profiles of 94 $CD8^+$ T cells, 94 $CD4^+$ T cells as well as the combined set of both T cells (188 profiles) and the methylation profiles assigned to the selected diagnostic categories in the training cohort of Capper *et al.*[25,36,47] Significance of hypomethylation and hypermethylation was computed separately, as we aimed for the identification of CpG sites consistently hypo-/hypermethylated between lymphocytes and tumors across all 85 diagnostic categories. The mean *p*-value for differential methylation over the 85 statistical tests was used to rank CpGs. Only those CpGs with a mean *p*-value < 0.05/428799 (Bonferroni correction) were retained for further analysis. This resulted in 373, 387, and 362 hypomethylated as well as 287, 217, and 243 hypermethylated CpGs between immune cells and tumor for $CD8^+$, $CD4^+$, and the mixed set of lymphocytes, respectively. As there was a large overlap between those CpG sites identified for $CD4^+$ and $CD8^+$ T cells, we next used immune cell profiles from the peripheral blood to identify specific CpG sites for each cell type.

Differential methylation analysis between immune cells from the peripheral blood

Methylation data from sorted blood cells using flow cytometric analyses of $CD4^+$, $CD8^+$ T cells, monocytes ($CD14^+$), NK cells ($CD56^+$), B cells ($CD19^+$) as well as neutrophils, eosinophils, (mixed) granulocytes and (mixed) peripheral blood mononuclear cells (PBMCs) from six patients were obtained from Reinius *et al.* via the R-package flow.sorted.blood.450k.[36,47] In a first step, to obtain specific CpGs for $CD8^+$ and $CD4^+$ T cells, the mean *p*-value of a paired one-sided Welch's *t*-test was computed between $CD4^+$/$CD8^+$ cells and the remaining populations of blood cells in a similar way as for the analysis of immune cells and tumor. Next, we selected those CpGs whose mean beta values were consistently higher/lower in $CD4^+$/$CD8^+$ than in all the remaining blood cell populations (threshold = 0.1). Differentially methylated sites were defined by the intersection of these sets resulting in 601 and 959 hypomethylated as well as 857 and 954 hypermethylated CpGs specific for $CD8^+$ and $CD4^+$ T cells, respectively. Specific CpG sites for lymphocytes were computed analogously comparing the $CD4^+$ and $CD8^+$ profiles with the $CD14^+$, neutrophils, eosinophils and granulocytes using an unpaired *t*-test resulting in 19393 hypomethylated and 30604 hypermethylated sites.

Definition of specific immune signature and dimensionality reduction

To obtain the final signatures of specifically differentially methylated CpGs, the intersection of the hypo/hypermethylated CpGs from the comparison of immune cells and tumor was built with the hypo/hypermethylated CpG sites comparing the studied immune cell type and the immune cells from the peripheral blood (monocytes ($CD14^+$), NK cells ($CD56^+$), B cells ($CD19^+$), neutrophils, and eosinophils). Next, hypo- and hypermethylated CpGs were combined. This resulted in 44, 30, and 197 specific CpGs for $CD8^+$ T cells (DIME-CD8), $CD4^+$ T cells (DIME-CD4) and tumor-infiltrating lymphocytes (DIME-TIL), respectively.

Based on the assumption that the main variance over the methylation of these specific CpGs is explained by the amount of infiltrating lymphocytes, we applied principal component

analysis (PCA) on these CpGs on the brain tumor training dataset for dimensionality reduction and defined the first principal component as a marker for the immune cell population. The amount of immune infiltration on a new dataset can then be estimated by applying the learned PCA representation to the new dataset. The orientation of the first PCA component was selected to correlate with the negative of the mean of the hypomethylated sites. The R code for the computation of DIMEimmune estimates is available as supplementary material.

Statistical and computational analyses, data visualization

DNA methylation-based diagnostic classification of brain tumors was executed with the Heidelberg classifier v11b6.[25] Gene ontology enrichment analyses for significant CpGs were performed with the *gometh* function.[32] Gene expression-based estimates for T cells and $CD8^+$ T cells were computed as previously described.[8] Briefly, well-established methods for TIL estimation[16,17] were adapted for the use with brain tumors by manually reviewing the signatures and excluding unspecific genes using correlation analyses. The MeTIL score was computed as described by Jeschke *et al.*[29] MethylCIBERSORT analysis was performed as implemented by Chakravarthy *et al.* using the brain tumor-specific signature matrix as published by Grabovska *et al.*[12,13] Gene expression-based/immunohistological TIL scores and methylation-based TIL scores were correlated with Spearman's rank correlation coefficient and robust linear regression was applied as implemented in the R package MASS.[38] Differences in estimates of TILs between different tumor entities were computed with the Kruskal-Wallis test. Proportional hazards modeling was performed with the R-package survival introducing immune scores as continuous variables.[35] *P*-values < 0.05 were considered statistically significant.

Results

Definition of specific signatures for tumor-infiltrating lymphocytes

The major principle of our approach is differential DNA methylation analysis between tumor profiles and lymphocytes as well as between different immune cell populations in the peripheral blood (Figure 1). As reference for the different tumor types, 2,706 samples from 85 different entities published by Capper *et al.* were used.[25] For immune cells, we used large cohorts of profiles from $CD4^+$ T cells and $CD8^+$ T cells, published by Tserel *et al.*[40] Applying different *t*-test statistics on brain tumors and immune cells resulted in lymphocyte-specific CpG sites (Selected CpGs 1, Figure 1b, see Methods). As there was considerable overlap between specific sites identified for $CD4^+$ T cells and $CD8^+$ T cells, we next used differential methylation as well as a thresholding analysis on the average methylation value between immune cell populations from the peripheral blood to identify specific CpG sites for $CD4^+$ and $CD8^+$ T cells (selected CpGs 2 + 3, Figure 1b). Furthermore, a mixed TIL signature was computed by performing the differential methylation analysis for the combined set of $CD4^+$ and $CD8^+$ T cells analogously (see methods). The final signature was obtained by intersecting the three selected groups of CpGs. Based on the assumption that the main variance in tumor profiles for the selected CpGs is based on the number of infiltrating lymphocytes, PCA was applied for dimensionality reduction on the selected CpGs in the tumor cohort. The first principal component was defined as an estimate for the tumor-infiltrating immune cells (Figure 1b). To validate our method, we compared our estimated score for immune cells with lymphocyte counts based on immunohistological images and signatures obtained from gene expression analysis. To this end, the PCA obtained in the training phase is applied to the methylation values of the independent test samples for the CpGs of the signature. Hence, the first component is defined as an estimate for the infiltration of the corresponding lymphocyte population (Figure 1c).

Analysis of computed signatures

We defined specific methylation signatures for $CD4^+$ (DIME-CD4) and $CD8^+$ T cells (DIME-CD8) as well as a mixed signature for tumor-infiltrating lymphocytes (DIME-TIL) (Figure 2a-c). These can be used to estimate the amount of infiltration of the corresponding cell type from bulk methylation profiles. Both, specific hypomethylated and specific hypermethylated CpGs exist in the obtained signatures. DIME-TIL contains 105 hypermethylated and 92 hypomethylated CpGs. DIME-CD4 has 4 hypermethylated and 26 hypomethylated CpGs, and DIME-CD8 has 17 hypermethylated and 27 hypomethylated CpGs. To get an overview of all three signatures and compare them to MeTIL, they have been applied to the flow-sorted blood dataset[47] (Figure 2d).

The methylation values divide the heatmaps into two different vertical areas. One area corresponds to the studied immune cells, and the other to the rest of the cell populations. In DIME-CD4 (Figure 2b), CpGs are the most hypo/hypermethylated for the $CD4^+$ T cells. This is also the case for DIME-CD8 and $CD8^+$ T cells. (Figure 2c). In DIME-TIL, both T cell types are differentially methylated as expected (Figure 2a). Our scores were compared with those from the previously published method by Jeschke *et al.* (MeTIL), which is also shown in Figure 2d. In contrast to our signatures, MeTIL results in highest scores for the $CD19^+$ B cells, followed by the other lymphocyte populations, showing a different weighting of the lymphocyte subpopulation during TIL estimation. All studied methods were able to differentiate between lymphocytes and $CD14^+$ cells as well as granulocytes.

Next, gene ontology enrichment was performed to identify pathways related to the lymphocyte-specific CpG signatures (Table 1). The most significantly enriched pathways for the DIME-TIL were "regulation of innate immune response" (p = 3.51e-04, FDR = 0.64) and "T cell activation" (p = 5.86e-04, FDR = 0.64). The most significantly enriched pathways for the DIME-CD4 were "regulation of defense response to virus by virus" (p = 1.82e-05, FDR = 0.04) and "positive regulation of interleukin-2 biosynthetic process" (p = 1.19e-04, FDR = 0.13). Finally, the most significantly enriched pathways for the

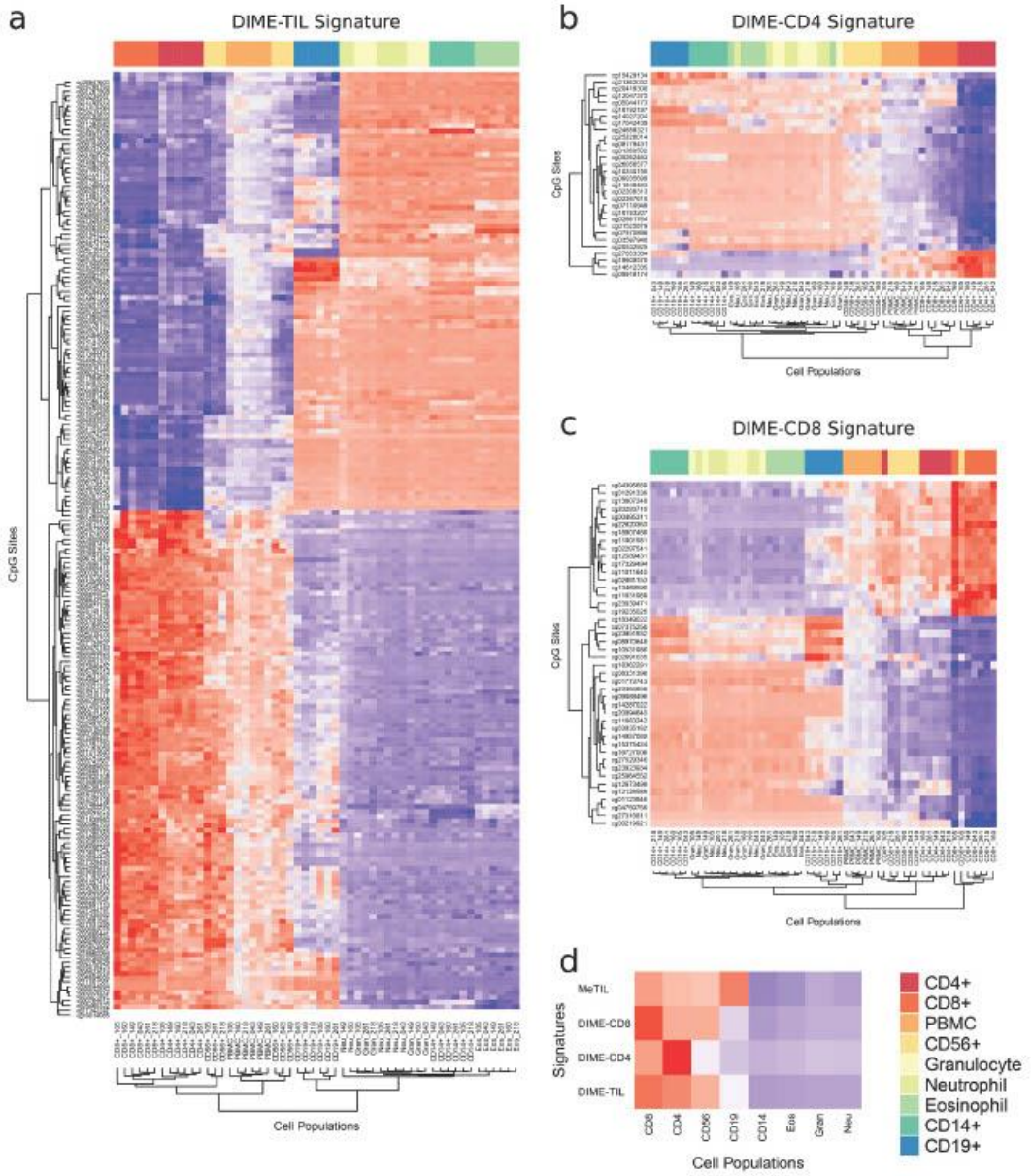


Figure 2. Visualization of the DIMEimmune signatures using heatmaps. Each colorful dot represents the methylation value of the corresponding CpG in the specified sample; a. DIME-TIL signature for tumor-infiltrating lymphocytes; b. DIME-CD4 signature of $CD4^+$ T cells; c. DIME-CD8 signature for $CD8^+$ T cells. d. Overview comparing estimates based on the three aforementioned signatures and MeTIL on the flow-sorted blood dataset. The DNA methylation values of samples of the same type have been averaged.

DIME-CD8 were "type I interferon signaling pathway" (p = 1.10e-04, FDR = 0.24) and "interferon-gamma-mediated signaling pathway" (p = 5.14e-04, FDR = 0.45). Most of the enrichment results are not significant after correction for multiple testing. Nonetheless, the most enriched categories are almost all related to the immune system, suggesting that the identified CpGs are related to genes belonging to immune system-related processes.

Table 1. Enrichment analysis of immune signatures. Results are tabulated for: DIME-TIL signature of tumor-infiltrating lymphocytes; DIME-CD4 signature of $CD4^+$ T cells; DIME-CD8 signature of $CD8^+$ T cells.

Cell Type	ID	TERM	DE	P.DE	FDR
TIL	GO:0045088	regulation of innate immune response	3	3.51e-04	0.64
	GO:0042110	T cell activation	4	5.86e-04	0.64
	GO:0060337	type I interferon signaling pathway	4	1.61e-03	1.00
	GO:0030217	T cell differentiation	3	4.00e-03	1.00
	GO:0031295	T cell costimulation	3	8.40e-03	1.00
	GO:1900017	positive regulation of cytokine production involved in inflammatory response	2	9.34e-03	1.00
	GO:0001816	cytokine production	2	9.54e-03	1.00
	GO:0030101	natural killer cell activation	2	9.86e-03	1.00
	GO:0043551	regulation of phosphatidylinositol 3-kinase activity	2	1.12e-02	1.00
	GO:0048535	lymph node development	2	1.16e-02	1.00
CD4	GO:0050690	regulation of defense response to virus by virus	3	1.82e-05	0.04
	GO:0045086	positive regulation of interleukin-2 biosynthetic process	2	1.19e-04	0.13
	GO:0006953	acute-phase response	2	4.76e-04	0.35
	GO:0050829	defense response to Gram-negative bacterium	2	1.20e-03	0.51
	GO:0033572	transferrin transport	2	1.27e-03	0.51
	GO:0034097	response to cytokine	2	1.56e-03	0.51
	GO:0042102	positive regulation of T cell proliferation	2	1.63e-03	0.51
	GO:0050731	positive regulation of peptidyl-tyrosine phosphorylation	2	7.97e-03	1.00
	GO:0033674	positive regulation of kinase activity	1	1.72e-02	1.00
	GO:0046039	GTP metabolic process	1	1.73e-02	1.00
CD8	GO:0060337	type I interferon signaling pathway	3	1.10e-04	0.24
	GO:0060333	interferon-gamma-mediated signaling pathway	3	5.14e-04	0.45
	GO:0043551	regulation of phosphatidylinositol 3-kinase activity	2	6.14e-04	0.45
	GO:0061512	protein localization to cilium	2	9.56e-04	0.52
	GO:0042110	T cell activation	2	3.09e-03	1.00
	GO:0048701	embryonic cranial skeleton morphogenesis	2	3.82e-03	1.00
	GO:0002479	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent	2	8.92e-03	1.00
	GO:0060071	Wnt signaling pathway, planar cell polarity pathway	2	1.59e-02	1.00
	GO:0019835	cytolysis	1	1.81e-02	1.00
	GO:0032897	negative regulation of viral transcription	1	1.97e-02	1.00

Validation and comparison of immune cell estimates with immunohistological results

As a first validation, methylation-based scores for infiltration of T cells (DIME-TIL, MeTIL, and the sum of the T cell signatures of MethylCIBERSORT) were compared to T cells counts from immunohistological images in a series of 47 diagnostic cases from our institution (Figure 3, Methods). Examples of cases with low/absent, intermediate, and high infiltration of $CD3^+$ cells are shown In Figure 3a-c. The DIME-TIL score showed the strongest correlation with the number of $CD3^+$ cells ($R = 0.74$, $p = 2e\text{-}09$), followed by MeTIL ($R = 0.32$, $p = 0.029$). The MethylCIBERSORT T cell signatures also showed a positive correlation with the number of $CD3^+$ cells quantified using immunohistochemistry ($R = 0.26$), but the correlation was not significant.

Validation and comparison of immune cell estimates with gene expression-based results

As a second benchmark, we used results from RNAseq and gene expression analysis. The validation data sets included 5 types of tumors: 763 medulloblastomas (MB), 129 ependymomas (EPN), 88 atypical teratoid/rhabdoid tumors (ATRT), 532 lower-grade gliomas (LGG), and 64 glioblastomas (GBM), for which both methylation and gene expression/RNAseq data were available. The remaining samples, with methylation data only, could not be used for this analysis. Our approach was also compared to the results for the two previously published methylation-based methods (MethylCIBERSORT and MeTIL). Whereas MeTIL provides only a global TIL score, MethylCIBERSORT can estimate several immune and stromal subpopulations. The comparative analysis for DIMEimmune and MeTIL is shown in Figure 4a.

The correlation of DIME-TIL with the gene expression-based T cell signature (Figure 4a) is best for GBM ($R = 0.72$, $p < 2.2e\text{-}16$), followed by LGG ($R = 0.56$, $p < 2.2e\text{-}16$), ATRT ($R = 0.53$, $p = 2.3e\text{-}6$), EPN ($R = 0.28$, $p = 0.0012$) and MB ($R = 0.13$, $p = 0.00032$). For MeTIL, the correlation is also best for GBM ($R = 0.6$, $p = 3.8e\text{-}7$), followed by LGG ($R = 0.5$, $p < 2.2e\text{-}16$), ATRT ($R = 0.42$, $p = 3e\text{-}04$) and EPN ($R = 0.21$, $p = 0.019$), whereas it is not significant for MB ($R = 0.042$, $p = 0.25$). Overall, correlation is superior for all 5 tumor entities for DIMEimmune based estimates compared to MeTIL based estimates.

Next, DIME-CD4 and CD4 estimates from MethylCIBERSORT ("CD4_Eff" + "Treg") were compared to the gene expression-based estimation of T cells (Figure 4b). For DIME-CD4, correlation is again best for GBM ($R = 0.5$, $p = 3.3e\text{-}5$), followed by LGG ($R = 0.42$, $p < 2.2e\text{-}16$) and ATRT ($R = 0.41$, $p = 0.00047$) and lowest for MB ($R = 0.072$, $p = 0.045$), whereas it is positive, but not significant for EPN ($R = 0.092$, $p = 0.3$). Correlation between MethylCIBERSORT CD4 and gene expression-based T cells estimates are best for ATRT ($R = 0.25$, $p = 0.036$) followed by LGG ($R = 0.19$, $p = 7.8e\text{-}06$) and not significant for GBM ($R = 0.19$, $p = 0.14$), MB ($R = -0.0034$, $p = 0.93$) and EPN ($R = 0.084$, $p = 0.35$).

Finally, methylation-based estimates for CD8 were compared with gene expression-based estimation of $CD8^+$ T cells (Figure 4c). DIME-CD8 showed highest correlation in GBM ($R = 0.52$, $p = 1.6e\text{-}5$), followed by ATRT ($R = 0.39$, $p = 0.00093$), LGG ($R = 0.33$,

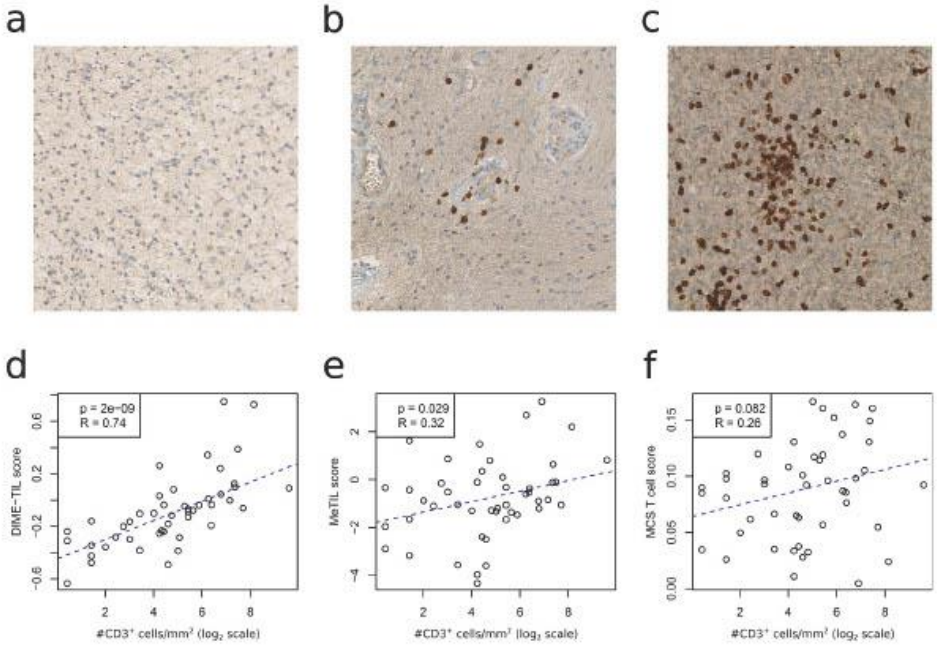


Figure 3. Immunohistological validation of methylation-based TIL estimates. Parallel immunohistological analyses for CD3 as well as methylation analyses were performed in a cohort of 47 brain tumors. Examples of tumors with low/absent $CD3^+$ cells; a. anaplastic astrocytoma IDH-mutant, WHO grade III, intermediate $CD3^+$ cells; b. rosette forming glioneuronal tumor, WHO grade I, and high numbers of $CD3^+$ cells; c. glioblastoma IDH wild-type, WHO grade IV. Immunohistological counts of $CD3^+$ cells are compared to DIME-TIL (d), MeTIL (e) and the sum of the T cells signatures analyzed by MethylCIBERSORT (f).

p = 3.7e-15) and EPN (R = 0.25, p = 0.0039), whereas it was not significant for MB (R = −0.071, p = 0.051). Correlation for MethylCIBERSORT CD8 estimates were best in ATRT (R = 0.46, p = 6.4e-5). Correlation in LGG (R = 0.085, p = 0.049) was weak and correlation in MB (R = −0.08, p = 0.027) was negative. Correlation in GBM (R = 0.18, p = 0.15) and EPN (R = 0.15, p = 0.085) was not significant.

Overall, our method resulted in an increased correlation with well-established gene expression-based results compared to previously published algorithms for all brain tumor types except for the MethylCIBERSORT estimate of $CD8^+$ T cells in ATRT.

Lymphocytic infiltration in brain tumors

Next, we used DIME-TIL estimation on the validation set of the Heidelberg Brain Tumor classifier (n = 1,104). Although larger, the training set was not used to avoid bias. As expected, the highest TIL score was found in CNS lymphoma (LYMPHO, Figure 5a). Further, there were high TIL scores in inflammatory and reactive control tissue (CONTR, REACT and CONTR, INFLAM). Other entities with high TIL scores included mesenchymal glioblastoma (GBM, MES), MYC ATRT, anaplastic pilocytic astrocytoma (ANA PA), melanoma (MELAN), chordoma (CHORDM) and pituitary adenomas (PITAD, TSH). Low TIL scores were found in posterior fossa B ependymoma (EPN, PF B), paraganglioma (PGG, nC) and group 4 medulloblastoma (MB, G4) (Figure 5a).

To get a more detailed insight into the distribution of TILs in different brain tumor subgroups, we applied our method to the validation data of LGG, GBM, MB, ATRT and EPN. Methylation-based diagnoses were computed with the Heidelberg Brain Tumor classifier.[25] As WHO grade III and IV tumors were present in the LGG and the GBM dataset, both datasets were combined to a glioma dataset. The differences in TIL scores between the 6 glioma entities were highly significant (Figure 5b, p = 1.6e-36). Estimated infiltration was lowest in IDH (Isocitrate dehydrogenase) mutated oligodendroglioma and in IDH mutated astrocytoma and highest in mesenchymal glioblastoma. For medulloblastoma, although statistically significantly different between molecular subgroups (p = 1.7e-49), the TIL score was overall low. The lowest number was found in Group 4 medulloblastoma and the largest number in Group 3 (Figure 5c). ATRTs showed prominent lymphocytic infiltration. The ATRT MYC subgroup had significantly more TILs than TYR (intermediate) and SHH (lowest) ATRT (p = 0.00012, Figure 5d). The TIL scores were significantly different between ependymoma subgroups (p = 4.4e-0.8, Figure 5e) with the lowest scores in the YAP and the PF B subgroup and the highest scores in the PF A and the RELA subgroup.

Associations of tumor-infiltrating lymphocytes with survival

Finally, we investigated the associations between the TIL estimates and overall survival in glioma, MB, ATRT and EPN. In the

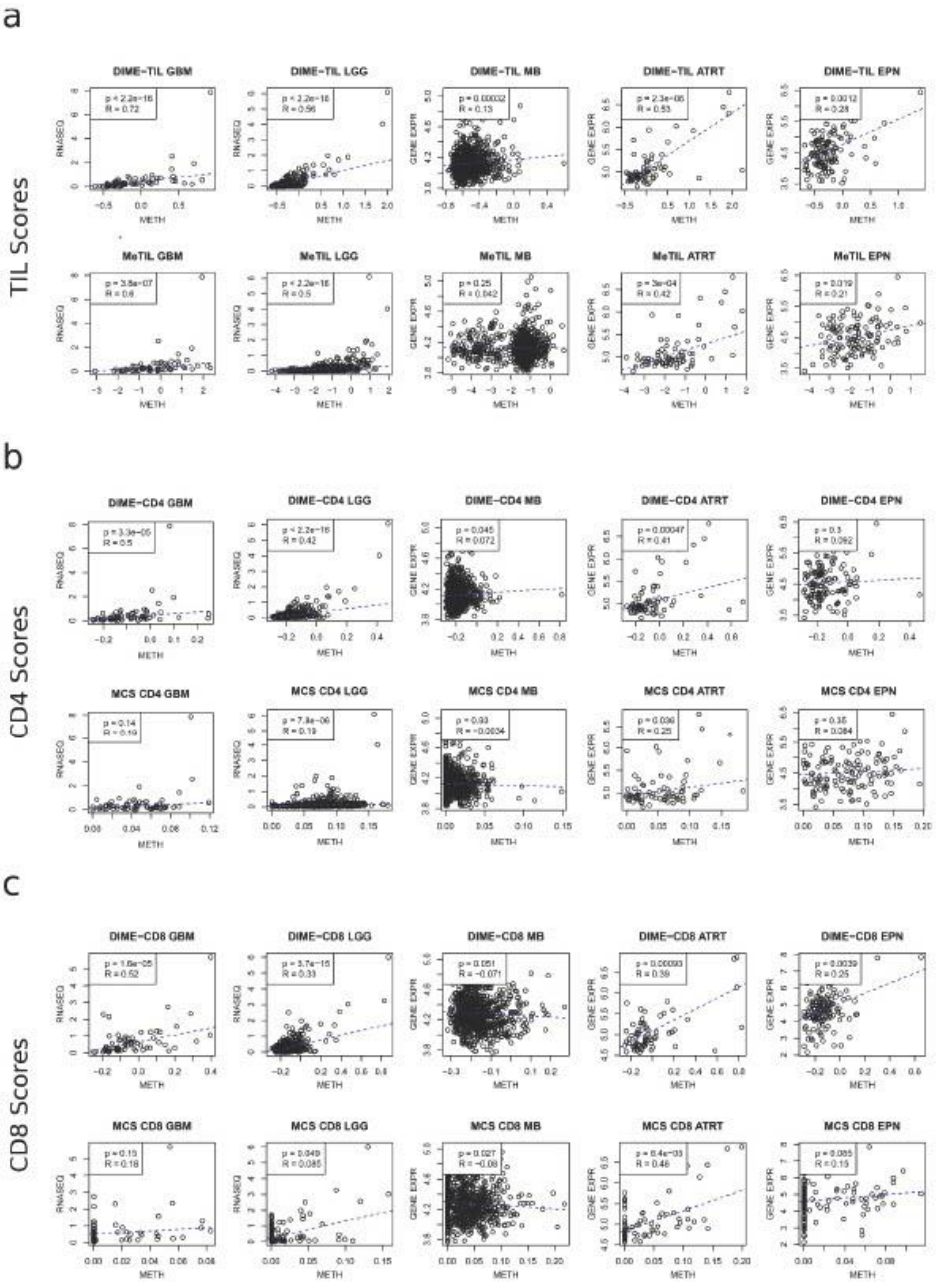


Figure 4. Comparison of DIMEImmune with previous studies for different tumor types (GBM, LGG, MB, ATRT, and EPN). Here, the gene expression/RNAseq-based results are used as the benchmark. In the diagrams, the dots are the samples, for which both gene expression/RNAseq and methylation data are available. The y-axis represents the gene expression/RNAseq-based estimates and the x-axis contains the methylation-based estimates; a. Comparison of DIME-TIL and MeTIL with gene expression/RNAseq-based estimation of T cells; b. Comparison of DIME-CD4 and $CD4^+$ T cells estimates obtained from MethylCIBERSORT (MCS) with gene expression/RNAseq-based estimation of $CD4^+$ T cells; c. Comparison of DIME-CD8 and $CD8^+$ T cells estimates obtained from MethylCIBERSORT with gene expression/RNAseq-based estimation of $CD8^+$ T cells.

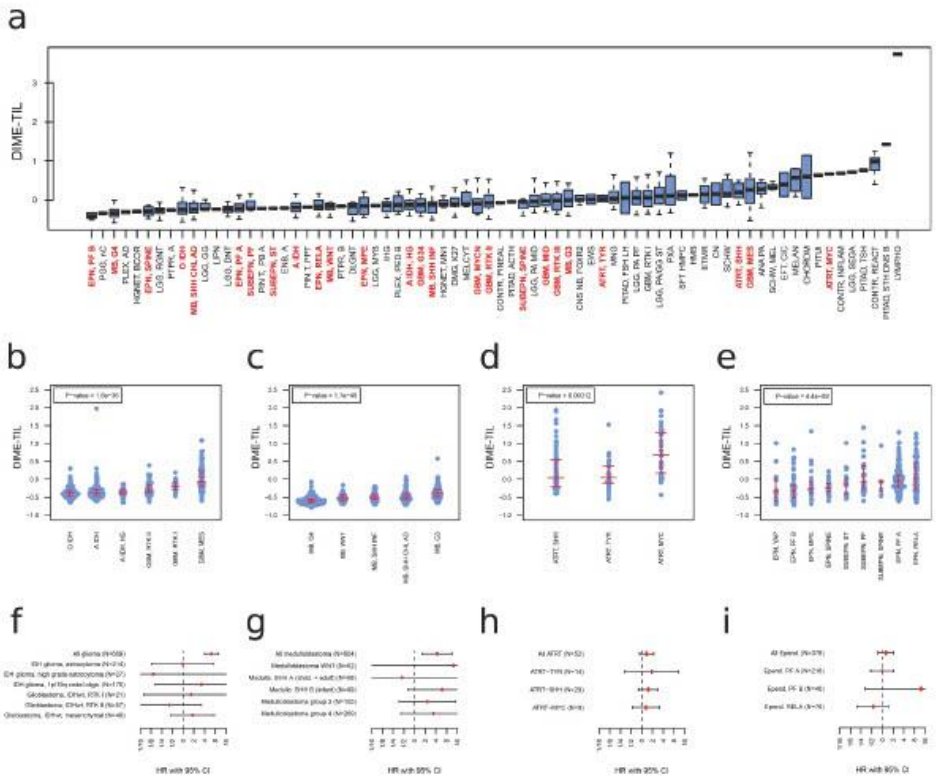


Figure 5. Clinical application of immune cell estimation. The estimated score of tumor-infiltrating lymphocytes for different subgroups of the Capper *et al.* validation data set (a), the TCGA glioma samples (b), medulloblastoma samples (c), ATRT (d), and ependymoma samples (e). The *p*-values are calculated based on the Kruskal-Wallis test. f-i.: Survival analysis for the prognostic relevance of DIME-TIL estimates for the same data as in panel b-e in the form of forest plots. HR: hazard ratio, CI: confidence interval.

overall glioma cohort, the DIME-TIL score was a strong negative prognostic factor (p = 6.06e-15, HR = 6.26). However, within methylation-based diagnostic subgroups, there was no significant association with survival (Figure 5f). Similarly, the DIME-TIL score was also a negative prognostic factor in medulloblastoma (p = 0.003, HR = 4.3), but not within methylation-defined subgroups (Figure 5g). There were no significant associations with survival in ATRT, likely due to low sample size compared to the other tumor types (Figure 5h). TIL scores were also not associated with survival in ependymoma and the studied subgroups, for which sufficient amount of survival data was available (Figure 5i).

Discussion

Although tumor-infiltrating lymphocytes in the CNS have been extensively studied using conventional immunohistological techniques (see Bienkowski and Preusser for review[5]) or gene expression-based analyses (e.g.[8,9,11,18,48]), reports using DNA methylation-based methods are still relatively rare. In contrast, DNA methylation has been central to brain tumor research in the last decade, and diagnostic tumor classification using DNA methylation data has been established in the diagnostic routine at several institutions.[49,50] This makes DNA methylation data widely available in the clinical routine in neuro-oncology. Although there are well-established techniques for immune cell estimation using, e.g., bulk gene expression data, we aimed for the development of a robust and user-friendly method, which can be used on DNA methylation data.

Specific reference signatures are key for quantification methods of immune cell signatures from DNA methylation data. Approaches based on MethylCIBERSORT do not only require specific signatures for immune cells, but also the tumor entity under consideration.[13] For the latter, DNA methylation data from cell lines are used.[12,13] While these data are widely available for common cranial and extra-cranial solid tumors, there are only few well-established cell lines for rare brain tumor entities. Further, *ex vivo* cell lines might have different methylation profiles from tumors *in vivo*. To allow for broader applicability of the method, we opted for differential methylation analysis using bulk data from brain tumors, which is far more widely available than cell lines.[25]

Overall, we established three methods to estimate $CD8^+$ T cells (DIME-CD8), $CD4^+$ T cells (DIME-CD4) and tumor-

infiltrating lymphocytes (DIME-TIL). Several steps of analysis were performed to obtain specific signatures for the cell type under consideration. Although several CpGs were identified as lymphocyte-specific, there were only few showing specific hypo/hypermethylation in $CD4^+$ and $CD8^+$ cells (Figure 2), underlining the difficulty to precisely separate lymphocyte subpopulations based on methylation data.

First, we used the correlation between the TIL scores of methylation-based methods (DIMEimmune, MeTIL and MethylCIBERSORT) and lymphocyte counts based on immunohistological images to validate and compare the efficiency of our method. Second, we used correlation analysis of TIL estimates from gene expression and methylation data as validation. These datatypes were selected as paired gene expression and methylation profiling is available for large cohorts of brain tumors. Overall, the DIME-TIL approach showed improved correlations with gene expression-based signatures compared to MeTIL and the implementation of MethylCIBERSORT, which was used for a pan-central nervous system cancer analysis by Grabovska *et al.*[12]

The MeTIL algorithm has been established using a methodologically related approach for immune cell estimation in breast cancer.[29] While it shows robust results for breast cancer, our approach gives better results for brain tumors. The correlation analysis for medulloblastoma in Figure 4a shows a bimodal distribution of the immune infiltration score, which is likely due to a nonspecific signature in medulloblastoma. We observed similar results in a previous gene expression-based study of immune infiltrates in medulloblastoma, where published lymphocyte signatures developed using extracranial tumors as reference contained genes linked to embryonal brain development.[9] This highlights the need for a brain tumor-specific TIL quantification algorithm, as algorithms and reference signatures developed for extracranial tumors might result in unspecific results in the CNS.

The MethylCIBERSORT algorithm showed mixed results in the correlation analysis with gene expression data and was outperformed by DIME-CD8 and DIME-CD4 for all comparisons except for the estimation of $CD8^+$ T cells in ATRT. For both $CD4^+$ and $CD8^+$ T cells the best performance of MethylCIBERSORT was in ATRT. The tumor reference signature used by Grabovska *et al.*[12] was composed solely of cell lines from medulloblastoma and rhabdoid tumors, but did not contain tumor cell lines from other brain tumors. This is likely to be the cause for better performance in ATRT than in gliomas, where the correlation is weak and highlights the need to include tumor-specific references. In their work, Grabovska *et al.* report immune cell contents from 75% to 100% for several brain tumors, in particular low-grade gliomas, which further underlines that tumor cells recognition by the algorithm is compromised for certain brain tumor entities due to unspecific tumor reference signatures. Further, the performance of CIBERSORT-based approaches might be compromised in tumors with low overall immune infiltration, as it has been previously shown for breast cancer.[19] Therefore, our approach might be more robust for the estimation of TIL abundance in CNS tumors, in particular for rare entities, for which there are only few or no cell line data available.

All algorithms showed only weak correlations with gene expression-based approaches in medulloblastoma, in particular for $CD8^+$ T cells. Only the correlation reported for DIME-TIL ($R = 0.13$, $p = 0.00032$) and DIME-CD4 ($R = 0.072$, $p = 0.045$) were significant. Medulloblastoma has rather low numbers for tumor-infiltrating immune cells as seen here (Figure 5a) as well as in other studies.[9,12] This might result in an unfavorable signal-to-noise ratio, which might be particularly pronounced for methylation analysis due to bimodal distribution of the underlying data. Therefore, immunohistology, gene expression (in particular RNA-sequencing) or single cell sequencing-based methods might be more suitable to study tumors with very low amounts of infiltrating immune cells. Nonetheless, methylation-based analysis keeps the advantage of a much broader application field.

Overall, estimates of TIL abundance based on DIME-TIL were in line with previously reported findings. The high scores in CNS lymphoma as well as reactive and inflammatory tissue can be regarded as a validation of our method. Medulloblastoma showed rather low scores for tumor-infiltrating immune cells as previously reported.[9,51] However, the methylation-based analysis did not find larger amounts of T cells in the SHH subgroups as identified by gene expression analysis,[9,52] possibly due to a high noise to signal ratio.

Lower TIL scores in lower-grade gliomas as well as in *IDH* mutated compared to *IDH* wild-type gliomas are in line with previously reported results.[53–55] Further, the mesenchymal subgroup is well known to be particularly highly infiltrated by TILs.[6,8] The larger estimates of TILs in MYC ATRTs compared to SHH and TYR ATRT has also been previously reported.[10,56] PF A ependymoma are among the subgroups of ependymoma with the largest estimated number of TILs, which is well compatible with the inflammatory phenotype that has been shown to be a key feature of this subtype.[11,57]

Survival analysis showed a significant negative prognostic effect of the DIME-TIL score in the overall glioma and the medulloblastoma cohort, but not within methylation subgroups. In glioma, this effect may be caused by increasing numbers of TILs in more aggressive WHO grade IV forms of glioma.[55] Similarly, the difference in TIL estimates in medulloblastoma subgroups, which are highest in the poor prognosis Group 3 medulloblastoma might also act as a confounder. Overall, there are conflicting reports in the literature on the prognostic effects of tumor-infiltrating immune cells in glioma and medulloblastoma.[6–9,12,51,55,58] These results emphasize the need for studies on the prognostic role of TILs in brain tumors in large cohorts of well-defined diagnostic subgroups, as the histological or molecular subtype might otherwise confound the results. To this end, methylation data-based methods offer a unique opportunity, as such data are very widely available.

To conclude, we established a method for the estimation of tumor-infiltrating lymphocytes in CNS tumors from methylation data based on differential analysis. The method showed better performance than previous methods taking lymphocyte counts based on immunohistological images and gene expression-based TIL estimates as reference. Estimates are most robust in tumors with more pronounced lymphocytic infiltrates. As it can be applied to any brain tumor entity, it can contribute to the identification of TILs as a prognostic or predictive factor in oncoming studies involving methylation data from tumors of the central nervous system.

List of abbreviations

CNS: central nervous system, TIL: tumor-infiltrating lymphocyte, DIMEimmune: Differential Methylation Analysis for Immune Cell Estimation, ATRT: atypical teratoid/rhabdoid tumors, MB: medulloblastoma, GBM: glioblastoma, LGG: lower-grade glioma, EPN: ependymoma, MCS: MethylCIBERSORT, HE: hematoxylin and eosin, GEO: Gene Expression Omnibus, PCA: principal component analysis, PBMC: peripheral blood mononuclear cell.

Declarations:

Ethics approval and consent to participate

Informed consent was obtained for all patients prior to methylation and immunohistological analyses.

Availability of data and material

The previously published data analyzed in this study are available on Gene Expression Omnibus and the European Genome-Phenome Archive (see methods). Raw methylation data generated within this study are available from the corresponding author upon reasonable request for non-commercial use. The code necessary to compute DIMEimmune estimates is available as supplemental material for non-commercial use. Any other generated code is available from the corresponding authors upon reasonable request for non-commercial use.

Competing interests

The authors declare no potential conflicts of interests.

Authors' contributions (CRediT)

Conceptualization: All authors, Methodology: SS+MB+US, Formal Analysis: SS+MB, Investigation: All authors, Resources: MB and US, Data Curation: SS+MB Writing – Original Draft: SS+MB, Writing – Review & Editing: All authors, Visualization: SS+MB+MM, Supervision: SS+MB, Funding: US.

Funding

SS, US, and MB were supported by the Fördergemeinschaft Kinderkrebszentrum Hamburg. MM was supported by the Hamburger Krebsgesellschaft and the Else Kröner-Fresenius-Stiftung.

References

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144:646–674. doi:10.1016/j.cell.2011.02.013.
2. Quail DF, Joyce JA. The Microenvironmental Landscape of Brain Tumors. Cancer Cell. 2017;31:326–341. doi:10.1016/j.ccell.2017.02.009.
3. Hendry S, Salgado R, Gevaert T, Russell PA, John T, Thapa B, Christie M, Vijver K V D, Estrada M V, Gonzalez-Ericsson P I, et al. Assessing Tumor-infiltrating Lymphocytes in Solid Tumors: a Practical Review for Pathologists and Proposal for a Standardized Method From the International Immunooncology Biomarkers Working Group: Part 1: Assessing the Host Immune Response, TILs in Invasive Breast Carcinoma and Ductal Carcinoma In Situ, Metastatic Tumor Deposits and Areas for Further Research. Adv Anat Pathol. 2017;24:235–251. doi:10.1097/PAP.0000000000000162.
4. Hendry S, Salgado R, Gevaert T, Russell PA, John T, Thapa B, Christie M, Vijver KVD, Estrada MV, Gonzalez-Ericsson P I, et al. Assessing Tumor-Infiltrating Lymphocytes in Solid Tumors: a Practical Review for Pathologists and Proposal for a Standardized Method from the International Immuno-Oncology Biomarkers Working Group: Part 2: TILs in Melanoma, Gastrointestinal Tract Carcinomas, Non-Small Cell Lung Carcinoma and Mesothelioma, Endometrial and Ovarian Carcinomas, Squamous Cell Carcinoma of the Head and Neck, Genitourinary Carcinomas, and Primary Brain Tumors. Adv Anat Pathol. 2017;24:311–335. doi:10.1097/PAP.0000000000000161.
5. Bieńkowski M, Preusser M. Prognostic role of tumour-infiltrating inflammatory cells in brain tumours: literature review. Curr Opin Neurol. 2015;28(6):647–658. doi:10.1097/WCO.0000000000000251.
6. Rutledge WC, Kong J, Gao J, Gutman DA, Cooper LAD, Appin C, Park Y, Scarpace L, Mikkelsen T, Cohen M L, et al. Tumor-infiltrating lymphocytes in glioblastoma are associated with specific genomic alterations and related to transcriptional class. Clin Cancer Res. 2013;19:4951–4960. doi:10.1158/1078-0432.CCR-13-0551.
7. Marinari E, Allard M, Gustave R, Widmer V, Philippin G, Merkler D, Tsantoulis P, Valérie D, Dietrich, P-Y. Inflammation and lymphocyte infiltration are associated with shorter survival in patients with high-grade glioma. Oncoimmunology. 2020;9:1779990. doi:10.1080/2162402X.2020.1779990.
8. Bockmayr M, Klauschen F, Maire CL, Rutkowski S, Westphal M, Lamszus K, Schüller U, Mohme M. Immunologic profiling of mutational and transcriptional subgroups in pediatric and adult high-grade gliomas. Cancer Immunol Res. 2019;7:1401–1411. doi:10.1158/2326-6066.CIR-18-0939.
9. Bockmayr M, Mohme M, Klauschen F, Winkler B, Budczies J, Rutkowski S, Schüller U. Subgroup-specific immune and stromal microenvironment in medulloblastoma. Oncoimmunology. 2018;7:e1462430. doi:10.1080/2162402X.2018.1462430.
10. Leruste A, Tosello J, Ramos RN, Tauziède-Espariat A, Brohard S, Han Z-Y, Beccaria K, Andrianteranagna M, Caudana P, Nikolic J, et al. Clonally expanded T cells reveal immunogenicity of rhabdoid tumors. Cancer Cell. 2019;36:597–612.e8. doi:10.1016/j.ccell.2019.10.008.
11. Griesinger AM, Josephson RJ, Donson AM, Mulcahy Levy JM, Amani V, Birks DK, Hoffman LM, Furtek SF, Reigan P, Handler MH, et al. Interleukin-6/STAT3 Pathway Signaling Drives an Inflammatory Phenotype in Group A Ependymoma. Cancer Immunol Res. 2015;3:1165–1174. doi:10.1158/2326-6066.CIR-15-0061.
12. Grabovska Y, Mackay A, O'Hare P, Crosier S, Finetti M, Schwalbe EC, Pickles JC, Fairchild AR, Avery A, Cockle J, et al. Pediatric pan-central nervous system tumor analysis of immune-cell infiltration identifies correlates of antitumor immunity. Nat Commun. 2020;11:4324. doi:10.1038/s41467-020-18070-y.
13. Chakravarthy A, Furness A, Joshi K, Ghorani E, Ford K, Ward MJ, King EM, Lechner M, Marafioti T, Quezada SA, et al. Pan-cancer deconvolution of tumour composition using DNA methylation. Nat Commun. 2018;9:3220. doi:10.1038/s41467-018-05570-1.
14. Klauschen F, Müller K-R, Binder A, Bockmayr M, Hägele M, Seegerer P, Wienerta S, Prunerie G, Mariaf S de, Badve S, et al. Scoring of tumor-infiltrating lymphocytes: from visual estimation to machine learning. Semin Cancer Biol. Elsevier, 2018;151–157. doi:10.1016/j.semcancer.2018.07.001.
15. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12:453–457. doi:10.1038/nmeth.3337.
16. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17:218. doi:10.1186/s13059-016-1070-5.
17. Danaher P, Warren S, Dennis L, D'Amico L, White A, Disis ML, Geller MA, Odunsi K, Beechem J, Fling SP. Gene expression markers of Tumor Infiltrating Leukocytes. J Immunother Cancer. 2017;5:18. doi:10.1186/s40425-017-0215-8.

18. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med. 2015;21:938–945. doi:10.1038/nm.3909.
19. Ali HR, Chlon L, Pharoah PDP, Markowetz F, Caldas C. Patterns of Immune Infiltration in Breast Cancer and Their Clinical Implications: a Gene-Expression-Based Retrospective Study. PLoS Med. 2016;13:e1002194. doi:10.1371/journal.pmed.1002194.
20. Cavalli FMG, Remke M, Rampasek L, Peacock J, Shih DJH, Luu B, Gröbner S, Segura-Wang M, Zichner T, Rudneva VA, et al. Intertumoral Heterogeneity within Medulloblastoma Subgroups. Cancer Cell. 2017;31(737–754.e6). doi:10.1016/j.ccell.2017.05.005.
21. Northcott PA, Buchhalter I, Morrissy AS, Hovestadt V, Weischenfeldt J, Ehrenberger T, et al. The whole-genome landscape of medulloblastoma subtypes. Nature. 2017;547:311–317. doi:10.1038/nature22973.
22. Pajtler KW, Witt H, Sill M, Jones DTW, Hovestadt V, Kratochwil F, Wani K, Tatevossian R, Punchihewa C, Johann P, et al. Molecular classification of ependymal tumors across all CNS compartments, histopathological grades, and age groups. Cancer Cell. 2015;27:728–743. doi:10.1016/j.ccell.2015.04.002.
23. Torchia J, Golbourn B, Feng S, Ho KC, Sin-Chan P, Vasiljevic A, Norman JD, Guilhamon P, Garzia L, Agamez NR, et al. Integrated (epi)-genomic analyses identify subgroup-specific therapeutic targets in CNS Rhabdoid Tumors. Cancer Cell. 2016;30:891–908. doi:10.1016/j.ccell.2016.11.003.
24. Sturm D, Orr BA, Toprak UH, Hovestadt V, Jones DTW, Capper D, Sill M, Buchhalter I, Northcott PA, Leis I, Ryzhova M, et al. New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs. Cell. 2016;164:1060–1072. doi:10.1016/j.cell.2016.01.015.
25. Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, Koelsche C, Sahm F, Chavez L, Reuss DE, et al. DNA methylation-based classification of central nervous system tumours. Nature. 2018;555:469–474. doi:10.1038/nature26000.
26. Jurmeister P, Bockmayr M, Seegerer P, Bockmayr T, Treue D, Montavon G, Vollbrecht C, Arnold A, Teichmann D, Bressem K, et al. Machine learning analysis of DNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases. Sci Transl Med. 2019;11: doi:10.1126/scitranslmed.aaw8513.
27. Orozco JIJ, Knijnenburg TA, Manughian-Peter AO, Salomon MP, Barkhoudarian G, Jalas JR, Wilmott JS, Hothi P, Wang X, Takasumi Y, et al. Epigenetic profiling for the molecular classification of metastatic brain tumors. Nat Commun. 2018;9:4627. doi:10.1038/s41467-018-06715-y.
28. Moran S, Martínez-Cardús A, Sayols S, Musulén E, Balañá C, Estival-Gonzalez A, Moutinho C, Heyn H, Diaz-Lagares A, Castro de Moura M, et al. Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. Lancet Oncol. 2016;17:1386–1395. doi:10.1016/S1470-2045(16)30297-2.
29. Jeschke J, Bizet M, Desmedt C, Calonne E, Dedeurwaerder S, Garaud S, Koch A, Larsimont D, Salgado R, Van den Eynden G, et al. DNA methylation-based immune response signature improves patient diagnosis in multiple cancers. J Clin Invest. 2017;127:3090–3102. doi:10.1172/JCI91095.
30. The R Core Team. R: a language and environment for statistical computing. 2019; Available from: https://www.r-project.org/
31. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA, Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30:1363–1369. doi:10.1093/bioinformatics/btu049.
32. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. Bioinformatics. 2016;32(2):286–288. doi:10.1093/bioinformatics/btv560.
33. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32(18):2847–2849. doi:10.1093/bioinformatics/btw313.
34. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. 2016;44:e71. doi:10.1093/nar/gkv1507.
35. Therneau TM, Grambsch PM Modeling Survival Data: extending the Cox Model [Internet]. Springer-Verlag:New York. 2000 [cited 2020 Sep 23]. Available from: https://www.springer.com/gp/book/9780387987842
36. FlowSorted.Blood.450k [Internet]. Bioconductor. [cited 2020 Sep 23]. Available from: http://bioconductor.org/packages/FlowSorted.Blood.450k/
37. Eklund A. beeswarm: the Bee Swarm Plot, an Alternative to Stripchart [Internet]. 2016 [cited 2020 Sep 23]. Available from: https://CRAN.R-project.org/package=beeswarm
38. Ripley B, Venables B, Bates DM, ca 1998) KH (partial port, ca 1998) AG (partial port, Firth D. MASS: support Functions and Datasets for Venables and Ripley's MASS [Internet]. 2020 [cited 2020 Sep 23]. Available from: https://CRAN.R-project.org/package=MASS
39. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinformatics. 2008;24(13):1547–1548. doi:10.1093/bioinformatics/btn224.
40. Tserel L, Kolde R, Limbach M, Tretyakov K, Kasela S, Kisand K, Saare M, Vilo J, Metspalu A, Milani L, et al. Age-related profiling of DNA methylation in CD8+ T cells reveals changes in immune response and transcriptional regulator genes. Sci Rep. 2015;5:13107. doi:10.1038/srep13107.
41. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30:207–210. doi:10.1093/nar/30.1.207.
42. Torchia J, Picard D, Lafay-Cousin L, Hawkins CE, Kim S-K, Letourneau L, Ra Y-S, Ching Ho K, Yu Chan TS, Sin-Chan P, et al. Molecular subgroups of atypical teratoid rhabdoid tumours in children: an integrated genomic and clinicopathological analysis. Lancet Oncol. 2015;16:569–582. doi:10.1016/S1470-2045(15)70114-2.
43. Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA, Morozova O, Yulia Newton Y, Radenbaugh A, Pagnotta SM, et al. Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. Cell. 2016;164:550–563. doi:10.1016/j.cell.2015.12.028.
44. Rahman M, Jackson LK, Johnson WE, Li DY, Bild AH, Piccolo SR. Alternative preprocessing of RNA-Sequencing data in The Cancer Genome Atlas leads to improved analysis results. Bioinformatics. 2015;31:3666–3672. doi:10.1093/bioinformatics/btv377.
45. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy—analysis of Affymetrix GeneChip data at the probe level. Bioinformatics. 2004;20:307–315. doi:10.1093/bioinformatics/btg405.
46. Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, Bunney WE, Myers RM, Speed TP, Akil H, et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Res. 2005;33:e175. doi:10.1093/nar/gni179.
47. Reinius LE, Acevedo N, Joerink M, Pershagen G, S-e D, Greco D, Söderhäll C, Scheynius A, Kere J. Differential DNA Methylation in Purified Human Blood Cells: implications for Cell Lineage and Studies on Disease Susceptibility. PLOS ONE. 2012;7:e41361. doi:10.1371/journal.pone.0041361.
48. Wang Q, Hu B, Hu X, Kim H, Squatrito M, Scarpace L, deCarvalho AC, Lyu S, Li P, Li Y, et al. Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. Cancer Cell. 2017;32(42–56.e6). doi:10.1016/j.ccell.2017.06.003.
49. Capper D, Stichel D, Sahm F, Jones DTW, Schrimpf D, Sill M, Schmid S, Hovestadt V, Reuss DE, Koelsche C, et al. Practical implementation of DNA methylation and copy-number-based CNS tumor diagnostics: the Heidelberg experience. Acta Neuropathol. 2018;136:181–210. doi:10.1007/s00401-018-1879-y.
50. Pickles JC, Stone TJ, Jacques TS. Methylation-based algorithms for diagnosis: experience from neuro-oncology. J Pathol. 2020;250(5):510–517. doi:10.1002/path.5397.

51. Vermeulen JF, Van Hecke W, Adriaansen EJM, Jansen MK, Bouma RG, Villacorta Hidalgo J, Fisch P, Broekhuizen R, Spliet WGM, Kool M, et al. Prognostic relevance of tumor-infiltrating lymphocytes and immune checkpoints in pediatric medulloblastoma. Oncoimmunology. 2017;7. Available from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5790383/.
52. Margol AS, Robison NJ, Gnanachandran J, Hung LT, Kennedy RJ, Vali M, Dhall G, Finlay JL, Erdreich-Epstein A, Krieger MD, et al. Tumor-associated macrophages in SHH subgroup of medulloblastomas. Clin Cancer Res. 2015;21:1457–1465. doi:10.1158/1078-0432.CCR-14-1144.
53. Berghoff AS, Kiesel B, Widhalm G, Wilhelm D, Rajky O, Kurscheid S, French PJ. Correlation of immune phenotype with IDH mutation in diffuse glioma. Neuro-oncology. 2017;19:1460–1468. doi:10.1093/neuonc/nox054.
54. Weenink B, Draaisma K, Ooi HZ, Kros JM, Pae SS, Debets R, et al. Low-grade glioma harbors few CD8 T cells, which is accompanied by decreased expression of chemo-attractants, not immunogenic antigens. Sci Rep. 2019;9:14643. doi:10.1038/s41598-019-51063-6.
55. Lohr J, Ratliff T, Huppertz A, Ge Y, Dictus C, Ahmadi R, Grau S, Hiraoka N, Eckstein V, Ecker RC, et al. Effector T-Cell Infiltration Positively Impacts Survival of Glioblastoma Patients and Is Impaired by Tumor-Derived TGF-β. Clin Cancer Res. 2011;17:4296–4308. doi:10.1158/1078-0432.CCR-10-2557.
56. Chun H-JE, Johann PD, Milne K, Zapatka M, Buellesbach A, Ishaque N, Iskar M, Erkek S, Wei L, Tessier-Cloutier B, et al. Identification and Analyses of Extra-Cranial and Cranial Rhabdoid Tumor Molecular Subgroups Reveal Tumors with Cytotoxic T Cell Infiltration. Cell Rep. 2019;29(2338–2354.e7). doi:10.1016/j.celrep.2019.10.013.
57. Hoffman LM, Donson AM, Nakachi I, Griesinger AM, Birks DK, Amani V, Hemenway MS, Liu AK, Wang M, Hankinson TC, et al. Molecular sub-group-specific immunophenotypic changes are associated with outcome in recurrent posterior fossa ependymoma. Acta Neuropathol. 2014;127:731–745. doi:10.1007/s00401-013-1212-8.
58. Han S, Zhang C, Li Q, Dong J, Liu Y, Huang Y, Jiang T, Wu A. Tumour-infiltrating CD4+ and CD8+ lymphocytes as predictors of clinical outcome in glioma. Br J Cancer. 2014;110:2560–2568. doi:10.1038/bjc.2014.162.

3. Abstract / Zusammenfassung

3.1. English

The interaction between central nervous system (CNS) tumors and infiltrating lymphocytes plays an important role in tumor initiation, propagation, and response to treatment. Traditionally, gene expression-based methods have been successfully utilized to study the tumor microenvironment. Recently, methylation data have been increasingly employed in molecular diagnostics. However, there are still few methods for analyzing the tumor microenvironment using this data type. To address this gap, we developed DIMEimmune (Differential Methylation Analysis for Immune Cell Estimation) based on differential methylation and principal component analysis. This tool estimates CD4+ and CD8+ T cells abundance, as well as tumor-infiltrating lymphocyte (TIL) scores from bulk methylation data.

DIMEimmune was validated by comparing its results to gene expression-based estimates and immunohistochemical counts of TIL. Across diverse tumor samples, DIMEimmune showed stronger correlations with both of these benchmarks than previously published methods such as MeTIL and MethylCIBERSORT. To show its prognostic impact, DIMEimmune was applied to another dataset of various tumor types. TIL estimates were low in medulloblastoma, high in ATRTs, particularly in the MYC subgroup, and higher for high-grade glioma than for low-grade glioma. DIMEimmune could show prognostic insights in the overall cohort of gliomas and medulloblastomas but not within methylation-based diagnostic subgroups.

In conclusion, DIMEimmune offers reliable TIL estimates and has the potential to be a prognostic or predictive tool in future CNS tumor studies.

3.2. Deutsch

Die Interaktion von Tumoren des zentralen Nervensystems (ZNS) mit infiltrierenden Lymphozyten wurde als ein wichtiger Faktor erkannt, der eine wesentliche Rolle bei der Tumorentstehung, -wachstum, und Therapieantwort spielt. Traditionell wurden Genexpressionsbasierte Methoden erfolgreich genutzt, um die Tumormikroumgebung zu untersuchen. In letzter Zeit werden zunehmend Methylierungsdaten in der molekularen Diagnostik eingesetzt. Es gibt jedoch immer noch nur wenige Methoden zur Analyse der Tumormikroumgebung anhand dieses Datentyps. Um diese Lücke zu schließen, wurde DIMEimmune (Differential Methylation Analysis for Immune Cell Estimation) entwickelt, die auf differenzieller Methylierung und Hauptkomponentenanalyse basiert. Dieses Werkzeug kann die Häufigkeit von CD4+ und CD8+ T-Zellen sowie die Werte von Tumor-infiltrierenden Lymphozyten (TIL) aus Methylierungsdaten schätzen.

DIMEimmune wurde validiert, indem seine Ergebnisse mit genexpressionsbasierten Schätzungen und immunhistochemischen Zählungen von TIL verglichen wurden. Die Korrelation zwischen den TIL-Werten von DIMEimmune und beiden dieser Referenzmethoden war besser als bei zuvor veröffentlichten Methoden, wie MeTIL und MethylCIBERSORT, in

verschiedenen Tumorproben. Um die prognostische Relevanz von DIMEimmune zu zeigen, wurde die Methode auf einen weiteren Datensatz mit verschiedenen Tumorarten angewendet. Dabei waren die TIL-Schätzungen bei Medulloblastomen niedrig, bei ATRT hoch, insbesondere in der MYC-Untergruppe, und bei hochgradigen Gliomen höher als bei niedriggradigen Gliomen. DIMEimmune konnte prognostische Erkenntnisse in der gesamten Kohorte von Gliomen und Medulloblastomen zeigen, jedoch nicht innerhalb der diagnostischen Untergruppen.

Zusammenfassend bietet DIMEimmune zuverlässige TIL-Schätzungen und könnte in zukünftigen Studien zu ZNS-Tumoren als prognostisches oder prädiktives Werkzeug eingesetzt werden.

4. References

1. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA, Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30:1363–1369. doi:10.1093/bioinformatics/btu049.
2. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17:218. doi:10.1186/s13059-016-1070-5.
3. Berghoff AS, Kiesel B, Widhalm G, Wilhelm D, Rajky O, Kurscheid S, French PJ. Correlation of immune phenotype with IDH mutation in diffuse glioma. 2017;19:1460–1468. doi:10.1093/neuonc/nox054. Neuro-oncology.
4. Bockmayr M, Mohme M, Klauschen F, Winkler B, Budczies J, Rutkowski S, Schüller U. Subgroup-specific immune and stromal microenvironment in medulloblastoma. Oncoimmunology. 2018;7:e1462430. doi:10.1080/2162402X.2018.1462430.
5. Bockmayr M, Harnisch K, Pohl LC, Schweizer L, Mohme T, Körner M, et al. Comprehensive profiling of myxopapillary ependymomas identifies a distinct molecular subtype with relapsing disease. Neuro-Oncology. 2022;24(10):1689–1699.
6. Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, Koelsche C, Sahm F, Chavez L, Reuss DE, et al. DNA methylation-based classification of central nervous system tumours. Nature. 2018;555:469–474. doi:10.1038/nature26000.
7. Capper D, Stichel D, Sahm F, Jones DTW, Schrimpf D, Sill M, Schmid S, Hovestadt V, Reuss DE, Koelsche C, et al. Practical implementation of DNA methylation and copy-number-based CNS tumor diagnostics: the Heidelberg experience. Acta Neuropathol. 2018;136:181–210. doi:10.1007/s00401-018-1879-y.
8. Cavalli FMG, Remke M, Rampasek L, Peacock J, Shih DJH, Luu B, Gröbner S, Segura-Wang M, Zichner T, Rudneva VA, et al. Intertumoral Heterogeneity within Medulloblastoma Subgroups. Cancer Cell. 2017;31(737–754.e6). doi:10.1016/j.ccell.2017.05.005.
9. Chakravarthy A, Furness A, Joshi K, Ghorani E, Ford K, Ward MJ, King EM, Lechner M, Marafioti T, Quezada SA, et al. Pan-cancer deconvolution of tumour composition using DNA methylation. Nat Commun. 2018;9:3220. doi:10.1038/s41467-018- 05570-1.
10. Chun H-JE, Johann PD, Milne K, Zapatka M, Buellesbach A, Ishaque N, Iskar M, Erkek S, Wei L, Tessier-Cloutier B, et al. Identification and Analyses of Extra-Cranial and Cranial Rhabdoid Tumor Molecular Subgroups Reveal Tumors with Cytotoxic T Cell Infiltration. Cell Rep. 2019;29(2338–2354.e7). doi:10.1016/j.celrep.2019.10.013.
11. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. 2016;44:e71. doi:10.1093/nar/gkv1507.

12. Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, Bunney WE, Myers RM, Speed TP, Akil H, et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Res. 2005;33:e175. doi:10.1093/nar/gni179.
13. Danaher P, Warren S, Dennis L, D'Amico L, White A, Disis ML, Geller MA, Odunsi K, Beechem J, Fling SP. Gene expression markers of Tumor Infiltrating Leukocytes. J Immunother Cancer. 2017;5:18. doi:10.1186/s40425-017-0215-8.
14. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinformatics. 2008;24(13):1547–1548. doi:10.1093/bioinformatics/btn224.
15. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30:207–210. doi:10.1093/nar/30.1.207.
16. Eklund A. beeswarm: the Bee Swarm Plot, an Alternative to Stripchart [Internet]. 2016 [cited 2020 Sep 23]. Available from: https://CRAN.R-project.org/package=beeswarm
17. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy—analysis of Affymetrix GeneChip data at the probe level. Bioinformatics. 2004;20:307–315. doi:10.1093/bioinformatics/btg405.
18. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med. 2015;21:938–945. doi:10.1038/nm.3909.
19. Grabovska Y, Mackay A, O'Hare P, Crosier S, Finetti M, Schwalbe EC, Pickles JC, Fairchild AR, Avery A, Cockle J, et al. Pediatric pan-central nervous system tumor analysis of immune-cell infiltration identifies correlates of antitumor immunity. Nat Commun. 2020;11:4324. doi:10.1038/s41467-020- 18070-y.
20. Griesinger AM, Josephson RJ, Donson AM, Mulcahy Levy JM, Amani V, Birks DK, Hoffman LM, Furtek SF, Reigan P, Handler MH, et al. Interleukin-6/STAT3 Pathway Signaling Drives an Inflammatory Phenotype in Group A Ependymoma. Cancer Immunol Res. 2015;3:1165–1174. doi:10.1158/2326-6066.CIR-15- 0061.
21. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32(18):2847–2849. doi:10.1093/bioinformatics/btw313.
22. Han S, Zhang C, Li Q, Dong J, Liu Y, Huang Y, Jiang T, Wu A. Tumour-infiltrating CD4+ and CD8+ lymphocytes as predictors of clinical outcome in glioma. Br J Cancer. 2014;110:2560–2568. doi:10.1038/bjc.2014.162.
23. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144:646–674. doi:10.1016/j.cell.2011.02.013.
24. Hendry S, Salgado R, Gevaert T, Russell PA, John T, Thapa B, Christie M, Vijver K V D, Estrada M V, Gonzalez-Ericsson P I, et al. Assessing Tumor-infiltrating Lymphocytes in Solid Tumors: a Practical Review for Pathologists and Proposal for a Standardized Method From the International Immunooncology Biomarkers Working Group: Part 1: Assessing the Host Immune Response, TILs in Invasive Breast Carcinoma and Ductal Carcinoma In Situ, Metastatic Tumor Deposits and Areas for Further Research. Adv Anat Pathol. 2017;24:235–251. doi:10.1097/PAP.0000000000000162.

25. Hendry S, Salgado R, Gevaert T, Russell PA, John T, Thapa B, Christie M, Vijver KVD, Estrada MV, Gonzalez-Ericsson P I, et al. Assessing Tumor-Infiltrating Lymphocytes in Solid Tumors: a Practical Review for Pathologists and Proposal for a Standardized Method from the International Immuno-Oncology Biomarkers Working Group: Part 2: TILs in Melanoma, Gastrointestinal Tract Carcinomas, Non-Small Cell Lung Carcinoma and Mesothelioma, Endometrial and Ovarian Carcinomas, Squamous Cell Carcinoma of the Head and Neck, Genitourinary Carcinomas, and Primary Brain Tumors. Adv Anat Pathol. 2017;24:311–335. doi:10.1097/PAP.0000000000000161.
26. Hoffman LM, Donson AM, Nakachi I, Griesinger AM, Birks DK, Amani V, Hemenway MS, Liu AK, Wang M, Hankinson TC, et al. Molecular sub-group-specific immunophenotypic changes are associated with outcome in recurrent posterior fossa ependymoma. Acta Neuropathol. 2014;127:731–745. doi:10.1007/s00401-013-1212-8.
27. Jaffe AE. FlowSorted.Blood.450k: Illumina HumanMethylation data on sorted blood cell populations [Internet]. Bioconductor. [cited 2020 Sep 23]. Available from: http://bioconductor.org/packages/FlowSorted.Blood.450k/
28. Jeschke J, Bizet M, Desmedt C, Calonne E, Dedeurwaerder S, Garaud S, Koch A, Larsimont D, Salgado R, Van den Eynden G, et al. DNA methylation-based immune response signature improves patient diagnosis in multiple cancers. J Clin Invest. 2017;127:3090–3102. doi:10.1172/JCI91095.
29. Johann PD, Altendorf L, Efremova EM, Holsten T, Steinbügl M, Nemes K, et al. Recurrent atypical teratoid/rhabdoid tumors (AT/RT) reveal discrete features of progression on histology, epigenetics, copy number profiling, and transcriptomics. Acta Neuropathol. 2023;146(4):527-41. doi: 10.1007/s00401-023-02608-7.
30. Jurmeister P, Bockmayr M, Seegerer P, Bockmayr T, Treue D, Montavon G, Vollbrecht C, Arnold A, Teichmann D, Bressem K, et al. Machine learning analysis of DNA methylation profiles distinguishes primary lung squamous cell carcinomas from head and neck metastases. Sci Transl Med. 2019;11: doi:10.1126/scitranslmed.aaw8513.
31. Klauschen F, Müller K-R, Binder A, Bockmayr M, Hägele M, Seegerer P, Wienerta S, Prunerie G, Mariaf S de, Badve S, et al. Scoring of tumor-infiltrating lymphocytes: from visual estimation to machine learning. Semin Cancer Biol. Elsevier, 2018;151–157. doi:10.1016/j.semcancer.2018.07.001.
32. Kresbach C, Dottermusch M, Eckhardt A, Ristow I, Paplomatas P, Altendorf L, et al. Atypical neurofibromas reveal distinct epigenetic features with proximity to benign peripheral nerve sheath tumor entities. Neuro Oncol. 2023 Sep;25(9):1644-1655. doi: 10.1093/neuonc/noad053.
33. Leruste A, Tosello J, Ramos RN, Tauzède-Espariat A, Brohard S, Han Z-Y, Beccaria K, Andrianteranagna M, Caudana P, Nikolic J, et al. Clonally expanded T cells reveal immunogenicity of rhabdoid tumors. Cancer ccell.2019.10.008.
34. Lohr J, Ratliff T, Huppertz A, Ge Y, Dictus C, Ahmadi R, Grau S, Hiraoka N, Eckstein V, Ecker RC, et al. Effector T-Cell Infiltration Positively Impacts Survival of Glioblastoma Patients and Is Impaired by Tumor-Derived TGF-β. Clin Cancer Res. 2011;17:4296–4308. doi:10.1158/1078-0432.CCR-10-2557.

35. Margol AS, Robison NJ, Gnanachandran J, Hung LT, Kennedy RJ, Vali M, Dhall G, Finlay JL, Erdreich-Epstein A, Krieger MD, et al. Tumor-associated macrophages medulloblastomas. Clin Cancer doi:10.1158/1078-0432.CCR-14-1144. in SHH subgroup of Res. 2015;21:1457–1465.
36. Marinari E, Allard M, Gustave R, Widmer V, Philippin G, Merkler D, Tsantoulis P, Valérie D, Dietrich, P-Y. Inflammation and lymphocyte infiltration are associated with shorter survival in patients with high-grade glioma. Oncoimmunology. 2020;9:1779990. doi:10.1080/2162402X.2020.1779990.
37. Moran S, Martínez-Cardús A, Sayols S, Musulén E, Balañá C, Estival-Gonzalez A, Moutinho C, Heyn H, Diaz-Lagares A, Castro de Moura M, et al. Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. Lancet Oncol. 2016;17:1386–1395. doi:10.1016/S1470-2045(16)30297-2.
38. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12:453–457. doi:10.1038/nmeth.3337.
39. Northcott PA, Buchhalter I, Morrissy AS, Hovestadt V, Weischenfeldt J, Ehrenberger T, et al. The whole-genome landscape of medulloblastoma subtypes. Nature. 2017;547:311–317. doi:10.1038/nature22973.
40. Orozco JIJ, Knijnenburg TA, Manughian-Peter AO, Salomon MP, Barkhoudarian G, Jalas JR, Wilmott JS, Hothi P, Wang X, Takasumi Y, et al. Epigenetic profiling for the molecular classification of metastatic brain tumors. Nat Commun. 2018;9:4627. doi:10.1038/s41467-018-06715-y.
41. Pajtler KW, Witt H, Sill M, Jones DTW, Hovestadt V, Kratochwil F, Wani K, Tatevossian R, Punchihewa C, Johann P, et al. Molecular classification of ependymal tumors across all CNS compartments, histopathological grades, and age groups. Cancer Cell. 2015;27:728–743. doi:10.1016/j.ccell.2015.04.002.
42. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. Bioinformatics. 2016;32(2):286–288. doi:10.1093/bioinformatics/btv560.
43. Pickles JC, Stone TJ, Jacques TS. Methylation-based algorithms for diagnosis: experience from neuro-oncology. J Pathol. 2020;250 (5):510–517. doi:10.1002/path.5397.
44. Quail DF, Joyce JA. The Microenvironmental Landscape of Brain Tumors. Cancer Cell. 2017;31:326–341. doi:10.1016/j.ccell.2017.02.009.
45. Reinius LE, Acevedo N, Joerink M, Pershagen G, S-e D, Greco D, Söderhäll C, Scheynius A, Kere J. Differential DNA Methylation in Purified Human Blood Cells: implications for Cell Lineage and Studies on Disease Susceptibility. PLOS ONE. 2012;7:e41361. doi:10.1371/journal.pone.0041361.
46. Ripley B, Venables B, Bates DM, ca 1998) KH (partial port, ca 1998) AG (partial port, Firth D. MASS: support Functions and Datasets for Venables and Ripley's MASS [Internet]. 2020 [cited 2020 Sep 23]. Available from: https://CRAN.R-project.org/package=MASS

47. Rutledge WC, Kong J, Gao J, Gutman DA, Cooper LAD, Appin C, Park Y, Scarpace L, Mikkelsen T, Cohen M L, et al. Tumor-infiltrating lymphocytes in glioblastoma are associated with specific genomic alterations and related to transcriptional class. Clin Cancer Res. 2013;19:4951–4960. doi:10.1158/1078-0432.CCR-13-0551.
48. Safaei S, Mohme M, Niesen J, Schüller U, Bockmayr M. DIMEimmune: Robust estimation of infiltrating lymphocytes in CNS tumors from DNA methylation profiles. OncoImmunology. 2021;10(1):1932365. doi:10.1080/2162402X.2021.1932365
49. Sturm D, Orr BA, Toprak UH, Hovestadt V, Jones DTW, Capper D, Sill M, Buchhalter I, Northcott PA, Leis I, Ryzhova M, et al. New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs. Cell. 2016;164:1060–1072. doi:10.1016/j.cell.2016.01.015.
50. The R Core Team. R: a language and environment for statistical computing. 2019; Available from: https://www.r-project.org/
51. Therneau TM, Grambsch PM Modeling Survival Data: extending the Cox Model [Internet]. Springer-Verlag:New York. 2000 [cited 2020 Sep 23]. Available from: https://www.springer.com/gp/book/9780387987842
52. Torchia J, Golbourn B, Feng S, Ho KC, Sin-Chan P, Vasiljevic A, Norman JD, Guilhamon P, Garzia L, Agamez NR, et al. Integrated (epi)-genomic analyses identify subgroup-specific therapeutic targets in CNS Rhabdoid Tumors. Cancer Cell. 2016;30:891–908. doi:10.1016/j.ccell.2016.11.003.
53. Torchia J, Picard D, Lafay-Cousin L, Hawkins CE, Kim S-K, Letourneau L, Ra Y-S, Ching Ho K, Yu Chan TS, Sin-Chan P, et al. Molecular subgroups of atypical teratoid rhabdoid tumours in children: an integrated genomic and clinicopathological analysis. Lancet Oncol. 2015;16:569–582. doi:10.1016/S1470-2045(15)70114-2.
54. Tserel L, Kolde R, Limbach M, Tretyakov K, Kasela S, Kisand K, Saare M, Vilo J, Metspalu A, Milani L, et al. Age-related profiling of DNA methylation in CD8+ T cells reveals changes in immune response and transcriptional regulator genes. Sci Rep. 2015;5:13107. doi:10.1038/srep13107.
55. Vermeulen JF, Van Hecke W, Adriaansen EJM, Jansen MK, Bouma RG, Villacorta Hidalgo J, Fisch P, Broekhuizen R, Spliet WGM, Kool M, et al. Prognostic relevance of tumor-infiltrating lymphocytes and immune checkpoints in pediatric medulloblastoma. Oncoimmunology. 2017;7. Available from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5790383/.
56. Wang Q, Hu B, Hu X, Kim H, Squatrito M, Scarpace L, deCarvalho AC, Lyu S, Li P, Li Y, et al. Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. Cancer Cell. 2017;32(42–56.e6). doi:10.1016/j.ccell.2017.06.003.
57. Weenink B, Draaisma K, Ooi HZ, Kros JM, Pae SS, Debets R, et al. Low-grade glioma harbors few CD8 T cells, which is accompanied by decreased expression of chemo-attractants, not immunogenic antigens. Sci Rep. 2019;9:14643. doi:10.1038/s41598-019-51063-6.

5. List of Abbreviations

ATRT	Atypical Teratoid/Rhabdoid Tumor
CD	Cluster of Differentiation
CI	Confidence Interval
CNS	Central Nervous System
CpG	Cytosine-phosphate-Guanine
DNA	Deoxyribonucleic Acid
FFPE	Formalin-Fixed Paraffin-Embedded
FDR	False Discovery Rate
FPKM	Fragments Per Kilobase Million
GBM	Glioblastoma Multiforme
GEO	Gene Expression Omnibus
GO	Gene Ontology
GTP	Guanosine Triphosphate
HE	Hematoxylin and Eosin
HR	Hazard Ratio
IDH	Isocitrate Dehydrogenase
LGG	Low-Grade Glioma
MCS	MethylCIBERSORT
MHC	Major Histocompatibility Complex
NK	Natural Killer Cells
NS	Not Significant
PCA	Principal Component Analysis
PBMCs	Peripheral Blood Mononuclear Cells
RNA	Ribonucleic Acid
RNAseq	RNA Sequencing
TAP	Transporter associated with Antigen Processing
TCGA	The Cancer Genome Atlas
TIL	Tumor Infiltrating Lymphocyte
Wnt	Wingless-related Integration Site

6. List of Figures

7. List of Tables

8. Declaration of the own contribution to the publication

I contributed to the conceptualization of the study and wrote the first draft of the manuscript. Specifically, I completed the following tasks:

- Table and Figures
    - Table 1
    - Figures 1, 2, 3d-f, and 4
- Materials and Method
    - Dataset acquisition and preprocessing of previously published methylation data
    - DNA methylation-based estimation of tumor-infiltrating lymphocytes
    - Differential methylation analysis between lymphocytes and CNS tumors
    - Differential methylation analysis between lymphocytes from the peripheral blood
    - Definition of T cell signatures and dimensionality reduction
- Results
    - Identification of specific signatures for tumor-infiltrating lymphocytes
    - Analysis of computed signatures
    - Validation and comparison of TIL estimates with immunohistological results
    - Validation and comparison of TIL estimates with gene expression-based results

9. Affirmation in lieu of an oath / Eidesstattliche Versicherung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe, insbesondere ohne entgeltliche Hilfe von Vermittlungs- und Beratungsdiensten, verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe. Das gilt insbesondere auch für alle Informationen aus Internetquellen.

Soweit beim Verfassen der Dissertation KI-basierte Tools („Chatbots“) verwendet wurden, versichere ich ausdrücklich, den daraus generierten Anteil deutlich kenntlich gemacht zu haben. Die „Stellungnahme des Präsidiums der Deutschen Forschungsgemeinschaft (DFG) zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG“ aus September 2023 wurde dabei beachtet.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich damit einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Datum: 21.02.2025 Unterschrift: 

10. Acknowledgements

Die Danksagung ist in zwei Teile untergliedert. Der erste Teil richtet sich an meine Familie, die wesentlich dazu beigetragen hat, dass ich überhaupt in der Lage bin, den Doktortitel anzustreben. Der zweite Teil richtet sich an meinen Betreuer und Doktorvater, die mir die Möglichkeit gegeben haben, am UCCH forschen zu dürfen.

Familie

Es wird gedankt:

- meinen Eltern, deren Liebe mich zur Welt gebracht hat. Sie haben alles, was in ihrer macht stand, investiert, damit ich das Beste im Leben erhalte.
- meinem Bruder und meinen Großeltern für ihre seelische Unterstützung.
- meiner Urgroßmutter, von der ich viel gelernt habe.
- Farid Assemi, der mich ständig unterstützt hat.

Betreuer und Doktorvater

Es wird gedankt:

- meinem Doktorvater, Prof. Schüller, der mich als Doktorand angenommen hat.
- meinem Betreuer, Prof. Bockmayr, der mich stets mit viel Geduld und Hingabe betreut und unterstützt hat.