

Compartment Mismatch Produces Near-Perfect Apparent Discrimination in Blood-Referenced DNA-Methylation Classifiers: A Public-Cohort Reanalysis of Hematologic-Malignancy Series

Frederic Scheer

fscheer@alercell.com

Alercell <https://orcid.org/0009-0000-9984-6337>

Research Article

Keywords: DNA methylation, cell composition, confounding, hematologic malignancy, public data reuse, deconvolution, laboratory-developed test, reproducibility

Posted Date: July 27th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10472780/v1>

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

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

Background.

Blood-based DNA-methylation classifiers are commonly validated by scoring publicly deposited disease cohorts against a healthy reference. Where the classifier is a deviation metric — the distance of a sample from a healthy baseline — it inherits the hematopoietic compartment of that baseline. Publicly deposited hematologic-malignancy series are heterogeneous in compartment: some are peripheral whole blood, others bone marrow, sorted fractions, or mobilized apheresis product. The consequence for validation has not been quantified.

Methods.

We defined a fully disclosed burden metric — mean absolute β -deviation from a healthy peripheral whole-blood reference (GSE40279, $n = 656$) — over 230,213 probes common and missingness-free across all cohorts. Compartment effect at zero disease was calibrated on GSE35069, in which six healthy donors were each profiled as whole blood, peripheral blood mononuclear cells (PBMC), granulocytes, and seven purified leukocyte subsets. Three hematologic-malignancy series were then scored against the same reference: GSE105420 (chronic myelomonocytic leukemia, CMML), GSE106600 (chronic myeloid leukemia, CML), and GSE118241 (myelofibrosis). Discrimination was quantified as the area under the receiver-operating-characteristic curve (AUC) with 10,000-fold bootstrap confidence intervals and Mann–Whitney tests. Leukocyte composition was estimated by robust partial correlation. Reference choice was varied across three independent healthy cohorts.

Results.

In healthy donors, compartment alone generated burden spanning 0.0229 (whole blood) to 0.0573 (CD8 + T cells), a range of 0.0344. Discrimination in the disease series tracked compartment matching rather than disease. Where cases and controls were drawn from different compartments, discrimination was near-perfect: AUC 1.000 (95% CI 1.000–1.000) for CML cells versus apheresis product, and 0.972 (0.931–1.000) for CMML bone marrow versus mobilized CD34 + peripheral blood. In the single compartment-matched comparison available — myelofibrosis versus control, both peripheral blood cells (GSE118241) — discrimination was absent: AUC 0.535 (0.316–0.763), $p = 0.83$. Holding disease constant and varying compartment within the same series produced substantially greater separation than disease did within matched compartment: AUC 0.827 (0.669–0.951), $p = 0.0016$. Estimated composition explained 45.2% of burden variance across the disease cohorts ($p = 6.7 \times 10^{-16}$); the mobilized CD34 + controls of GSE105420 carried a median neutrophil fraction of 0.909 against 0.450 in the marrow cases they were compared with. Diagnosis and compartment were completely collinear in GSE105420. Reference choice alone shifted whole-blood burden by 82.7% across three healthy cohorts and moved the metric 39% as much as compartment did.

Conclusions.

Compartment mismatch between reference and test cohort is sufficient to produce near-perfect apparent discrimination in the absence of any demonstrated disease effect. Blood-referenced classifiers validated on publicly deposited hematologic series without compartment matching, reported composition estimates, and a stated reference are not interpretable as evidence of disease discrimination. We propose three reporting requirements.

1. Introduction

A large and growing class of molecular diagnostics scores an individual's blood methylome against a healthy reference and treats the distance as evidence of disease. The construction is attractive: it requires no disease-specific training set, it produces a single interpretable scalar, and it can be evaluated on data already in the public domain. Its validation, in practice, almost always follows the same route — compute the metric on a publicly deposited disease series, compute it on a healthy series, and report the separation.

That route has an unexamined dependency. A deviation metric is defined relative to a reference, and it therefore inherits the reference's tissue and cellular composition. Bulk methylation is a composition-weighted average across the cell types present in a sample [1, 2], so two samples drawn from different hematopoietic compartments will differ in bulk methylation whether or not either is diseased. If the reference is peripheral whole blood and the test cohort is bone marrow, the metric registers that difference as burden. Nothing in the summary statistics typically reported — sensitivity, specificity, area under the curve — distinguishes that contribution from disease.

The exposure is neither hypothetical nor rare. Publicly deposited hematologic-malignancy methylation series are heterogeneous in compartment by design: studies of leukemic stem-cell biology profile sorted CD34 + fractions, studies of marrow pathology profile aspirates, and healthy comparators are frequently drawn from whatever material a center had available, including granulocyte colony-stimulating factor (G-CSF)-mobilized apheresis product. These are appropriate choices for the primary studies. They become a problem on reuse, when a secondary analysis treats a deposited matrix as interchangeable with peripheral blood.

This study asks one question: **when a blood-referenced deviation metric is applied to publicly deposited hematologic-malignancy cohorts, how much of the resulting separation is attributable to hematopoietic compartment rather than to disease?**

Answering it requires a calibrator in which compartment varies and disease does not. Reinius et al. profiled the same six healthy male donors as whole blood, PBMC, granulocytes, and seven purified leukocyte subsets on a single platform [3]. Scoring those samples against a whole-blood reference measures burden generated by compartment alone, with disease held at zero by construction. It also

requires at least one disease series in which cases and controls share a compartment, so that a matched comparison is possible. GSE118241 supplies one.

We use a fully disclosed metric rather than any proprietary panel, so that every value reported here is reproducible from public data by any group. The framing question — what cross-sectional public data can and cannot license about methylation-based hematologic surveillance — was set out previously [4]; this study supplies a quantitative answer to one part of it.

2. Methods

2.1 Burden metric

For a sample i with β -values β_p at probe p , and a healthy reference with per-probe mean β , burden is the mean absolute deviation

$$\mathbf{MB}_i = (1/|P|) \cdot \sum_{p \in P} |\beta_p - \beta|$$

over probe set P . Regulatory entropy is the normalized Shannon entropy of the per-sample deviation distribution. No proprietary or withheld panel is used anywhere in this study; both metrics are fully specified above and computable from the deposited matrices.

2.2 Probe space

All seven series were assayed on the Illumina HumanMethylation450 BeadChip (GPL13534). Deposited matrices carried 408,148 to 485,577 probes. P was defined as the intersection of probes present and free of missing values in every cohort, yielding **230,213 probes**. Analyses were run at full probe resolution; no subsampling was applied. A secondary probe set P_{var} , the 10,000 most variable probes in the primary reference, was retained for sensitivity analysis.

2.3 Reference cohorts

The primary reference is GSE40279 (peripheral whole blood, $n = 656$) [5]. Reference-choice sensitivity was assessed by repeating the compartment calibration against GSE51057 (buffy coat, $n = 329$) and GSE69270 (buffy coat, $n = 184$).

2.4 Compartment calibrator

GSE35069 [3] provides six healthy donors each profiled in ten compartments (whole blood, PBMC, granulocytes, neutrophils, eosinophils, CD14 + monocytes, CD19 + B cells, CD4 + T cells, CD8 + T cells, CD56 + NK cells; $n = 60$ arrays). Because all donors are healthy, burden computed on these samples is attributable to compartment alone.

2.5 Disease cohorts

GSE105420 (CMML) [6], GSE106600 (CML) [7], and GSE118241 (myelofibrosis) [8]. Tissue provenance was verified for every series against the primary publication and against the `source_name_ch1` and `characteristics_ch1` fields of the GEO series matrix (Table 1). Cohorts were classified as compartment-matched if cases and controls within the series shared a `source_name_ch1` value, and compartment-mismatched otherwise.

2.6 Deconvolution

Leukocyte proportions were estimated by robust partial correlation against a validated blood reference signature [9]. Of 333 reference probes, 208 (62.5%) survived the missingness-free intersection defining P ; estimates are therefore degraded and are reported with that caveat (§ 5). Composition models were fitted on the disease cohorts only, excluding both the calibrator and the healthy reference arm — the latter is scored against a reference constructed from itself, has burden near zero by construction, and anchors any pooled fit at the origin.

2.7 Statistics

Discrimination is reported as AUC with percentile bootstrap 95% confidence intervals (10,000 resamples) and two-sided Mann–Whitney U tests. Median differences carry bootstrap confidence intervals by the same procedure. Composition models are ordinary least squares of burden on the seven estimated fractions, reported as R^2 , adjusted R^2 , and the overall F-test p -value. We do not report the ratio of compartment effect to disease effect: the denominator is near zero in the matched comparison and the ratio is unstable under resampling (bootstrap median 2.0, interquartile range – 3.5 to 5.0). The two effects are reported separately instead. Analyses were performed in R 4.5.3 with EpiDISH 2.26.0, GEOquery 2.78.0, and matrixStats 1.5.0; the full session record accompanies the deposited code.

3. Results

3.1 Verified provenance

Table 1

Cohorts analyzed, with tissue provenance verified against primary publications and GEO series-matrix metadata. All series are GPL13534.

Accession	Role	Compartment(s)	Arrays
GSE40279	Primary healthy reference	Peripheral whole blood	656
GSE51057	Alternate reference	Buffy coat	329
GSE69270	Alternate reference	Buffy coat	184
GSE35069	Compartment calibrator	Whole blood, PBMC, granulocytes, 7 purified subsets	60
GSE105420	CMML	Bone marrow total cells (cases); CD34+-enriched mobilized peripheral blood (controls)	76
GSE106600	CML	CP-CML patient cells (cases); apheresis product (controls)	23
GSE118241	Myelofibrosis	Peripheral blood cells (cases and controls); granulocytes; bone marrow	45

Two of three disease series are compartment-mismatched between cases and controls. GSE118241 alone contains cases and controls in a shared compartment.

3.2 Compartment alone generates substantial burden in healthy donors

Scoring the calibrator against the whole-blood reference produced a monotonic ordering across compartments in donors with no disease (Table 2, Fig. 1). Burden ranged from 0.0229 in whole blood — the compartment of the reference itself — to 0.0573 in CD8 + T cells, a span of **0.0344**.

Table 2
Burden against the healthy whole-blood reference (GSE40279) by compartment, in six healthy donors profiled in all ten compartments (GSE35069). Median (interquartile range).

Compartment	<i>n</i>	Median	Q1	Q3
Whole blood	6	0.0229	0.0209	0.0252
PBMC	6	0.0320	0.0296	0.0348
Neutrophils	6	0.0331	0.0316	0.0335
Granulocytes	6	0.0333	0.0327	0.0352
CD14 + monocytes	6	0.0373	0.0361	0.0384
Eosinophils	6	0.0404	0.0395	0.0412
CD19 + B cells	6	0.0464	0.0460	0.0580
CD4 + T cells	6	0.0470	0.0463	0.0479
CD56 + NK cells	6	0.0499	0.0432	0.0644
CD8 + T cells	6	0.0573	0.0540	0.0608

Figure 1. Burden against a healthy whole-blood reference by hematopoietic compartment, six healthy donors, ten compartments each. Disease is absent by construction; all separation shown is compartment effect.

3.3 Discrimination tracks compartment matching, not disease

Burden by cohort and group is given in Table 3. The discrimination those values support depends entirely on whether the groups being compared share a compartment (Table 4).

Table 3
Burden by cohort and group. Median (interquartile range).

Cohort	Group	Compartment	<i>n</i>	Median	Q1	Q3
GSE105420	Healthy	CD34 + mobilized PB	10	0.0617	0.0592	0.0634
GSE105420	CMML	Bone marrow	65	0.0819	0.0761	0.0875
GSE106600	Normal	Apheresis product	11	0.0420	0.0402	0.0479
GSE106600	CML	CP-CML patient cells	12	0.0581	0.0553	0.0601
GSE118241	Control	Peripheral blood	6	0.0436	0.0401	0.0460
GSE118241	Myelofibrosis	Peripheral blood	19	0.0449	0.0375	0.0503
GSE118241	Myelofibrosis	Bone marrow	6	0.0525	0.0484	0.0565
GSE118241	Myelofibrosis	Granulocytes	14	0.0536	0.0508	0.0592

Table 4

Discrimination by compartment status. AUC with percentile bootstrap 95% confidence interval; median difference with bootstrap 95% confidence interval; two-sided Mann–Whitney *p*.

Comparison	Compartment status	AUC (95% CI)	Median Δ (95% CI)	<i>p</i>
CML cells vs apheresis product (GSE106600)	Mismatched	1.000 (1.000–1.000)	+ 0.0161 (+ 0.0080, + 0.0195)	5.6 × 10 ⁻⁵
CMML marrow vs mobilized CD34+ (GSE105420)	Mismatched	0.972 (0.931–1.000)	+ 0.0202 (+ 0.0162, + 0.0247)	1.8 × 10 ⁻⁶
Myelofibrosis vs control, peripheral blood (GSE118241)	Matched	0.535 (0.316–0.763)	+ 0.0013 (– 0.0061, + 0.0076)	0.83
Myelofibrosis PB vs myelofibrosis granulocytes (GSE118241)	<i>Disease held constant</i>	0.827 (0.669–0.951)	+ 0.0088 (+ 0.0035, + 0.0179)	0.0016
Myelofibrosis PB vs myelofibrosis bone marrow (GSE118241)	<i>Disease held constant</i>	0.763 (0.518–0.947)	+ 0.0076 (– 0.0012, + 0.0174)	0.059

Three observations follow.

First, where cases and controls occupy different compartments, the metric separates them essentially perfectly. In GSE106600 the separation is complete — every CML sample exceeds every apheresis-product control. In GSE105420 it is near-complete.

Second, in the only comparison where cases and controls share a compartment, no discrimination is detected. The point estimate of 0.535 is indistinguishable from chance and the median difference of + 0.0013 is two orders of magnitude smaller than the burden values themselves.

Third, and most directly, holding disease constant and varying compartment within a single series produces greater separation (AUC 0.827) than varying disease within a matched compartment (AUC 0.535). Compartment is the stronger discriminator of the two, measured in the same patients on the same platform.

The compartment-matched comparison rests on six controls, and its confidence interval reaches 0.763; a moderate true effect is not excluded (§ 5). The claim supported is not that myelofibrosis produces no burden signal in peripheral blood, but that the matched comparison does not reproduce anything approaching the near-perfect discrimination the mismatched comparisons report.

3.4 Composition accounts for a large share of burden, and differs systematically between mismatched groups

Estimated composition explained **45.2%** of burden variance across the disease cohorts (adjusted $R^2 = 0.428$, $n = 144$, F-test $p = 6.7 \times 10^{-16}$). Per cohort: GSE105420 $R^2 = 0.342$ ($p = 4.4 \times 10^{-5}$), GSE106600 $R^2 = 0.480$ ($p = 0.070$), GSE118241 $R^2 = 0.370$ ($p = 0.0052$).

The mechanism is visible directly in the estimated fractions (Table 5). In GSE105420, cases and controls differ by roughly half their neutrophil content — 0.450 against 0.909 — with correspondingly large differences in B, T, and NK fractions. That is not a disease contrast; it is a tissue contrast.

Table 5
Median estimated leukocyte fractions by cohort and group.

Cohort	Group	Neutrophil	CD4 + T	CD8 + T	Monocyte	B	NK
GSE105420	Mobilized CD34+ (controls)	0.909	0.003	0.002	0.079	0.003	0.007
GSE105420	Bone marrow (CMML)	0.450	0.044	0.057	0.197	0.125	0.093
GSE106600	Apheresis product	0.563	0.000	0.000	0.189	0.033	0.019
GSE106600	CP-CML patient cells	0.508	0.039	0.028	0.121	0.112	0.082
GSE118241	Peripheral blood	0.784	0.064	0.037	0.094	0.012	0.046
GSE118241	Granulocytes	0.971	0.000	0.000	0.020	0.002	0.001
GSE118241	Bone marrow	0.796	0.023	0.007	0.096	0.013	0.048

Figure 2. Observed burden against burden predicted from estimated composition alone, by cohort.

3.5 Diagnosis and compartment are completely collinear in GSE105420

Table 6

Diagnosis by sample source in GSE105420, from the GEO series matrix.

Sample source	CMML	Healthy	Unlabeled
Bone marrow, total cells	65	0	1
CD34+-enriched mobilized peripheral blood	0	10	0

No cell of this table permits compartment and diagnosis to be distinguished. When a compositional variable is perfectly confounded with the biological variable of interest, no adjustment recovers the separation, because no residual variation remains on which to condition — the same identifiability failure documented for batch correction under confounded designs [10, 11]. The AUC of 0.972 in Table 4 is therefore not an estimate of disease discrimination degraded by confounding; it is a quantity with no disease interpretation available at all.

3.6 Reference choice moves the metric substantially

Repeating the compartment calibration against two further healthy cohorts (Table 7) changed burden for identical samples by a median of 0.0133 across references — **39% of the entire compartment span**. Whole blood shifted most, from 0.0229 against GSE40279 to 0.0418 against GSE69270, an increase of **82.7%**. Reference choice also compressed resolution: the compartment span narrowed from 0.0344 under GSE40279 to 0.0241 under GSE69270.

Table 7
Median compartment burden under three independent healthy references.

Compartment	GSE40279	GSE51057	GSE69270	Range	$\Delta\%$ (69270 vs 40279)
Whole blood	0.0229	0.0281	0.0418	0.0189	+ 82.7
PBMC	0.0320	0.0342	0.0415	0.0095	+ 29.7
Neutrophils	0.0331	0.0381	0.0502	0.0171	+ 51.6
Granulocytes	0.0333	0.0382	0.0509	0.0176	+ 52.8
CD14 + monocytes	0.0373	0.0397	0.0504	0.0131	+ 35.2
Eosinophils	0.0404	0.0458	0.0570	0.0166	+ 41.2
CD19 + B cells	0.0464	0.0486	0.0549	0.0085	+ 18.3
CD4 + T cells	0.0470	0.0471	0.0537	0.0067	+ 14.3
CD56 + NK cells	0.0499	0.0550	0.0633	0.0134	+ 26.8
CD8 + T cells	0.0573	0.0610	0.0656	0.0082	+ 14.4

A burden value is therefore uninterpretable without both the compartment of the sample and the identity of the reference. Neither is conventionally reported.

4. Discussion

The central result is that compartment mismatch is *sufficient* to produce near-perfect apparent discrimination. Two public hematologic-malignancy series in which cases and controls occupy different hematopoietic compartments yield AUCs of 1.000 and 0.972 for a metric that, in the one series permitting a matched comparison, discriminates at 0.535. The same metric, the same disease family, the same array platform. Only the design differs.

That bulk methylation depends on cell composition is well established [1, 2]. What this analysis adds is the magnitude of the consequence for a validation practice that is common, cheap, and superficially rigorous — and the demonstration that the resulting statistics are not merely inflated but, in the collinear case, devoid of disease interpretation.

Three implications follow.

Compartment matching is a design requirement, not a caveat. A deviation metric calibrated on whole blood is not defined on marrow, on sorted fractions, or on apheresis product. Applying it across that boundary does not degrade the estimate; it changes what the estimate measures. The remedy is to match compartment at the design stage. Where no matched public series exists for a disease of interest, the correct conclusion is that public data cannot yet validate the classifier for that disease — not that a mismatched comparison will do in the interim.

Composition and reference must be reported alongside discrimination. Estimated cell proportions are computable from any deposited matrix at negligible cost [9, 12]. A discrimination metric reported without them, and without the identity of the reference cohort, cannot be evaluated by a reader: composition accounted for 45% of burden variance here, and reference choice alone moved the metric 39% as much as compartment did. We propose three minimum reporting requirements for blood-referenced classifiers validated on public data: (i) the compartment of every cohort, cases and controls separately; (ii) estimated composition, with discrimination reported before and after adjustment; (iii) the reference cohort, with a sensitivity analysis across at least one alternative.

Collinear designs should be identified and excluded, not adjusted. GSE105420 is a sound design for its primary purpose and an unusable one for secondary validation, because case and control compartments do not overlap at all. The defect is visible in a single cross-tabulation of two metadata fields (Table 6) and is invisible in every statistic computed downstream. A provenance check of this kind costs minutes and belongs before analysis, not after.

Implications for laboratory-developed tests. Analytical validity — whether an assay measures what it claims, reproducibly — is distinct from clinical validity, whether the measurement tracks disease. A blood-referenced burden metric validated across compartments has established neither: it has established that two tissues differ. For a test intended to operate on peripheral blood in a screening context, validation on marrow-derived or apheresis-derived cohorts supplies no support for the intended use, and any review that examines specimen type will identify this.

What this study does not establish. It does not show that methylation burden lacks disease signal in myeloid malignancy. The matched comparison rests on six controls and its confidence interval extends to 0.763; a real but moderate effect would not have been detected. What it shows is that the public cohorts most often used to demonstrate such signal cannot separate it from compartment, and that the near-perfect discrimination those cohorts produce is fully consistent with compartment effect alone. Establishing a genuine effect requires a compartment-matched study adequately powered for the effect size the matched data here are consistent with — considerably smaller than the mismatched cohorts imply.

5. Limitations

The compartment calibrator comprises six healthy male donors [3]. It establishes the scale of the compartment effect but not its variability across sex, age, or clinical state, and the values in Table 2 should be read as a calibration rather than as population parameters.

The compartment-matched comparison in GSE118241 has six controls against nineteen cases. It is underpowered for modest effects, and the negative result is properly read as the absence of large-effect discrimination rather than the demonstration of none.

Only 208 of 333 reference probes for the deconvolution signature survived the missingness-free intersection across seven cohorts, so composition estimates are degraded and the R^2 values in § 3.4 are best treated as lower bounds on the compositional contribution. Bone marrow additionally contains populations absent from any peripheral-blood deconvolution reference — nucleated erythroid precursors, immature granulocytic forms, megakaryocytes — so composition estimates for marrow samples are extrapolations and are reported as such.

The intersection reduced the probe space from a per-cohort range of 408,148–485,577 to 230,213. The burden metric is a mean over probes and is not systematically biased by this reduction, but cohort-specific probe attrition was not itself analyzed.

GSE105420 lacks a chronological-age field, preventing evaluation of age as a competing explanation in that cohort. Finally, this analysis covers three disease series; a systematic provenance audit across the wider set of hematologic methylation series in frequent secondary reuse would establish how widespread the exposure is, and is the natural extension of this work.

6. Conclusion

Hematopoietic compartment mismatch between reference and test cohort is sufficient to generate near-perfect apparent discrimination by a blood-referenced DNA-methylation burden metric, in the absence of any demonstrated disease effect. Across three public hematologic-malignancy series, discrimination tracked compartment matching rather than diagnosis: AUC 1.000 and 0.972 where compartments differed, 0.535 where they matched, and 0.827 for compartment alone with disease held constant. Estimated composition explained 45% of burden variance, and reference choice moved the metric 39% as much as compartment did. Blood-referenced classifiers validated on publicly deposited disease cohorts without compartment matching, reported composition, and a stated reference should not be read as evidence of disease discrimination.

7. Data availability

All data analyzed are publicly deposited in the NCBI Gene Expression Omnibus and were used as archived: GSE40279, GSE51057, GSE69270 (healthy references); GSE35069 (compartment calibrator); GSE105420, GSE106600, GSE118241 (disease series). Analysis code, the derived tables underlying every figure and table above, and the full session record are available at: [10.5281/zenodo.21517032](https://doi.org/10.5281/zenodo.21517032)

Declarations

Competing interests.

The author is Founder, Chief Executive Officer, and Chief Scientific Officer of Alercell, Inc., which is developing DNA-methylation-based diagnostics for hematologic malignancy. The findings reported here

constrain the interpretation of public-cohort validation for blood-referenced methylation classifiers generally, including approaches under development at Alercell. No proprietary panel, data, or method is used in this study; the metric is fully disclosed in § 2.1 and all data are public.

Use of artificial intelligence in manuscript preparation.

The author used Anthropic Claude during study design, analysis scripting, drafting, copy-editing, and verification of cited GEO accession provenance. All analyses were executed by the author and all findings were confirmed against primary publications and repository metadata. The author reviewed and revised the full manuscript, checked the scientific claims and references, and accepts responsibility for the final text. No AI system is listed as an author, consistent with COPE guidance on AI-assisted writing.

Funding.

Alercell, Inc. No external funding was received.

Author contributions.

F.S. designed the study, performed the analyses, and wrote the manuscript.

References

1. Houseman EA, Accomando WP, Koestler DC et al (2012) DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 13:86
2. Jaffe AE, Irizarry RA (2014) Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol* 15(2):R31
3. Reinius LE, Acevedo N, Joerink M et al (2012) Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS ONE* 7(7):e41361
4. Scheer F (2026) Static DNA-methylation burden discriminates hematologic-malignancy subtypes but does not license per-patient trajectory inference: a public-cohort analysis and an identifiability framework. *Research Square*. 10.21203/rs.3.rs-10397076/v1. Preprint
5. Hannum G, Guinney J, Zhao L et al (2013) Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell* 49:359–367
6. Palomo L, Malinverni R, Cabezón M et al (2018) DNA methylation profile in chronic myelomonocytic leukemia associates with distinct clinical, biological and genetic features. *Epigenetics* 13(1):8–18
7. Maupetit-Méhouas S, Court F, Bourgne C et al (2018) DNA methylation profiling reveals a pathological signature that contributes to transcriptional defects of CD34 + CD15– cells in early chronic-phase chronic myeloid leukemia. *Mol Oncol* 12(6):814–829

8. Martínez-Calle N, Pascual M, Ordoñez R et al (2019) Epigenomic profiling of myelofibrosis. *Haematologica* 104(8):1572
9. Teschendorff AE, Breeze CE, Zheng SC, Beck S (2017) A comparison of reference-based algorithms for correcting cell-type heterogeneity in epigenome-wide association studies. *BMC Bioinformatics* 18(1):105
10. Nygaard V, Rødland EA, Hovig E (2016) Methods that remove batch effects while retaining group differences may lead to exaggerated confidence in downstream analyses. *Biostatistics* 17(1):29–39
11. Zindler T, Frieling H, Neyazi A, Bleich S, Friedel E (2020) Simulating ComBat: how batch correction can lead to trouble when the batch effect is confounded with a biological variable. *BMC Genomics* 21(1):271
12. Salas LA, Zhang Z, Koestler DC et al (2022) Enhanced cell deconvolution of peripheral blood using DNA methylation for high-resolution immune profiling. *Nat Commun* 13(1):761

Figures

Figures 1 and 2 are available in the supplementary files section.

Supplementary Files

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

- [CompartmentMismatchProducesGraphicsanmdTablesavailability.pdf](#)