Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Leukocyte Composition Change Accounts for Most Apparent Ten-Year DNA Methylation Drift in Blood: A Two-Cohort Longitudinal Analysis

Frederic Scheer

fscheer@alercell.com

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

Research Article

Keywords: immunosenescence. cell-type deconvolution, leukocyte heterogeneity, biomarker validation, Shannon entropy,

Posted Date: August 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10733012/v1

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

Additional Declarations: The authors declare potential competing interests as follows: The author is the founder, chief executive officer, and chief scientific officer of Alercell, Inc., a company developing DNA methylation-based diagnostics for hematologic malignancy. The findings reported here constrain the interpretation of blood-based DNA methylation measures generally, including approaches under development at Alercell.

Abstract

Longitudinal DNA methylation measurement is increasingly proposed as a surveillance tool, on the premise that within-subject differencing removes the between-person and technical offsets that confound cross-sectional comparison. We tested whether that premise holds when the confounding variable is itself time-varying. We quantified change in a 48-locus, blood-referenced methylation burden score in two independent longitudinal cohorts of older adults — 86 paired whole-blood samples from the Danish Longitudinal Study of Aging Twins (120-month interval, Illumina 450K) and 16 paired buffy-coat samples from the Health, Aging and Body Composition study (60-month interval, Illumina EPIC) — and then adjusted each participant's change for concurrent change in leukocyte composition estimated by reference-based deconvolution. Unadjusted burden increased significantly in the Danish cohort (+0.000518 units/month; 95% CI, +0.000353 to +0.000682; $P = 2.2 \times 10^{-8}$), and the result survived correction for twin non-independence. After composition adjustment, 32.2% of that velocity remained and it was no longer distinguishable from zero (+0.000167 units/month; P = .106). Composition change alone explained 36.4% and 63.9% of the variance in burden change in the two cohorts, with concordant lymphoid-to-myeloid direction. Replicate arrays on physically distinct chips placed the technical standard deviation of a burden difference at 0.0328, indicating that 88% of the variance in 10-year change is biological rather than measurement error. Per-locus change ran toward $\beta = 0.5$ rather than along the panel's directional axis, and differential storage duration was excluded on two independent grounds. Within-subject differencing does not control leukocyte composition, because composition drifts alongside the outcome.

Introduction

Blood is the accessible tissue on which most human epigenetic biomarkers are built. Epigenetic age estimators, inflammation indices, and an expanding set of proposed cancer-surveillance assays all read DNA methylation from peripheral blood and compare it against a reference derived from other people (1-4). The confound that dominates this literature is cellular: whole blood is a mixture, the mixture proportions vary between individuals, and methylation differs sharply between leukocyte lineages (5-7). A cross-sectional difference between two people can therefore reflect the composition of their blood rather than any per-cell epigenetic state.

The standard remedy for a between-person confounder is a within-person design. If each participant is measured twice and the analysis operates on the difference, any offset that is constant within that participant — genotype, habitual cell composition, the plate on which they were run — cancels exactly. This reasoning is sound, and it is the explicit justification for the longitudinal designs now being proposed for methylation-based surveillance, in which an individual is monitored over time against their own baseline rather than against a population reference.

The reasoning holds, however, only for confounders that are time-invariant. Leukocyte composition in older adults is not. Immunosenescence produces a systematic contraction of the naive T-cell

compartment and a relative expansion of the myeloid compartment over the same decades in which methylation is being tracked (8-10). A confounder that changes in a consistent direction across participants does not cancel under differencing; it is differenced into the outcome. Whether this matters quantitatively — whether it accounts for a negligible fraction of observed longitudinal change or for most of it — is an empirical question that, to our knowledge, has not been answered directly for a locus-panel readout.

This concern is not new. The investigators who generated one of the two cohorts analyzed here reported that changes in cellular composition explain much of both the cross-sectional and the longitudinal variation in CpG methylation in their sample (12). Their observation was made genome-wide and was qualitative; what has been missing is a quantitative partition — how much of a measured longitudinal velocity survives adjustment, whether the partition replicates across cohorts and platforms, and whether the surviving remainder can be distinguished from measurement error. Those are the questions we take up here. We addressed that question in two independent cohorts of older adults with repeat blood sampling, using a 48-locus blood-referenced burden score as the test statistic. The score is one instance of a general construction — count or weight the deviations of a specimen from a healthy blood reference across a fixed locus set — that underlies many proposed methylation biomarkers, so the finding is not specific to this panel. We first measured unadjusted change, then adjusted it for concurrent change in deconvolution-estimated cell fractions, and then addressed three alternative explanations for the residual: measurement error, differential specimen storage, and the possibility that the change is directional along the panel's own axis rather than an undirected increase in disorder.

We report that composition change accounts for roughly two thirds of the apparent 10-year drift, that the composition-independent remainder cannot be distinguished from zero at this sample size, and that the change that does occur is entropic rather than directional. We also report the boundaries of that conclusion, which are wide: this is a failure to demonstrate composition-independent drift, not a demonstration of its absence.

Method

Cohorts and specimens

Two publicly deposited longitudinal cohorts met the requirement of repeat DNA methylation measurement on the same participants with full coverage of the locus panel.

The Danish Longitudinal Study of Aging Twins (LSADT) contributed 180 arrays from 43 twin pairs sampled in 1997 and again in 2007, a 120-month interval, with participants aged 73 to 82 years at intake (11). Specimens were whole blood assayed on the Illumina HumanMethylation450 BeadChip. After pairing on participant identifier, 86 individuals contributed two timepoints each. The deposit is unfiltered, retaining all 485 577 probes.

The Health, Aging and Body Composition (Health ABC) substudy contributed 37 arrays from 20 participants sampled at study years 1, 2, and 6, with a mean baseline age in the eighth decade (12). Specimens were buffy coat assayed on the Illumina MethylationEPIC BeadChip. Seventeen participants contributed two timepoints; 16 of these spanned the standard 60-month interval and one spanned 12 months. Because follow-up intervals were heterogeneous, the primary Health ABC analysis models elapsed time per participant as a covariate rather than dividing by a fixed divisor; a 60-month-only sensitivity arm is reported alongside it.

Both cohorts are de-identified public deposits. No institutional review board approval was required for secondary analysis of these data, and none was sought.

Burden score and reference panel

For each locus the observed β value is expressed as its deviation from a healthy whole-blood reference mean, divided by a fixed scale constant of 0.1 β units common to all 48 loci, rectified to the panel's stated direction for that locus so that only deviation in the specified direction contributes, and the resulting values are combined as a weighted mean with locus weights ranging from 0.166 to 0.542. The scale constant is not a per-locus empirical standard deviation; loci are therefore differentiated by the supplied weights alone and not by their variability in healthy blood. Because the constant is common to every locus and the rectification is positively homogeneous, it scales the burden score uniformly and affects no ratio, coefficient of determination, or test statistic reported here.

The healthy reference means are the medians of a published healthy whole-blood series (13). This is a material limitation and is treated as such in the Discussion: the reference cohort also contributed to the original selection of these loci, so absolute burden values are not independent of panel derivation. The present analysis is nonetheless insulated from that circularity, because every quantity reported here is a within-participant difference, and any reference offset that is constant across a participant's two timepoints cancels in the difference.

Panel identity was enforced computationally. The scoring routine refuses any locus set that does not match the canonical 48-probe list by identity and order, and refuses any specimen with incomplete coverage. Both cohorts returned 48 of 48. As a provenance check, the scoring engine reproduced a previously published healthy whole-blood burden median of 0.174 on an independent purified-fraction series to four decimal places (0.1736).

Cell-composition estimation

Cell fractions were estimated by reference-based deconvolution. The reference was constructed from a purified-leukocyte methylation series (6) covering seven populations — CD4+ T cells, CD8+ T cells, natural killer cells, B cells, monocytes, neutrophils, and eosinophils — by selecting the 60 most discriminating probes per direction per population from all 485 577 probes, yielding 713 reference probes. LSADT carried 713 of 713; Health ABC, on a different array platform, carried 666 of 713 (93.4%).

Deconvolution and burden scoring use disjoint probe sets, so the adjustment is not a regression of the outcome on a transformation of itself.

Statistical analysis

For each participant we computed Δburden = burden(t_2) − burden(t_1) and the elapsed interval Δt. Unadjusted velocity is the coefficient on elapsed time in a model of Δburden, equivalent in the homogeneous-interval case to the mean of Δburden/Δt.

Composition-adjusted velocity is the intercept of a linear model of Δburden on the vector of Δ cell fractions with elapsed time entered as above. Because estimated fractions sum to one at each timepoint, the seven Δproportion columns sum to zero and the design matrix is rank deficient. The primary specification therefore drops one population as the compositional reference; results are reported with eosinophils as reference, and the rank-deficient minimum-norm fit is reported alongside to confirm that the estimable quantities — intercept, fitted values, and R^2 — are identical (they were, to five decimal places). Individual cell-type coefficients from the rank-deficient fit are not interpretable and are not reported.

Because LSADT participants are twins, the nominal standard error is anticonservative. We report a worst-case correction treating each twin pair as a single independent unit (effective n = 43).

Retained velocity is expressed as adjusted divided by unadjusted magnitude. Where unadjusted velocity is not significantly different from zero, this ratio is a quotient of two near-zero quantities and is not interpretable; the analysis refuses to report a verdict in that case.

Technical variance from replicates

LSADT contains eight same-participant, same-year replicate array pairs, distributed across four participants at both timepoints. All eight pairs sit on physically distinct BeadChips, so the resulting estimate incorporates chip-to-chip variation and is not merely within-chip precision. The technical standard deviation of a burden difference was estimated from the mean absolute difference within replicate pairs as $\sigma_{technical} = \text{mean}|\Delta| \times \sqrt{(\pi/2)}$. Biological standard deviation was obtained by subtracting technical from total variance.

Storage-degradation and directionality tests

Baseline specimens were stored longer than follow-up specimens in both cohorts by construction. Progressive degradation of the longer-stored aliquot would drive β values away from 0.5 in the follow-up sample relative to baseline, producing a positive coefficient when per-locus mean Δβ is regressed on locus centrality, defined as the reference mean minus 0.5. It would also produce a dose-response, since the storage gap is 10 years in LSADT and 5 years in Health ABC. Both predictions were tested. Centrality was computed from the external reference mean rather than from the cohort's own baseline; regressing

on the observed baseline manufactures a negative coefficient through regression to the mean, an artifact that reached r = −0.71 on simulated null data.

The same model carries the panel's hyper/hypo direction vector as a second predictor, distinguishing undirected movement toward β = 0.5 from movement along the panel's directional axis. Because loci that are hypermethylated in the target condition are typically unmethylated in health, these predictors are correlated; we report their correlation and variance inflation factor, and a boundary-sensitivity analysis excluding loci whose reference mean lies near the β floor or ceiling. Panel Shannon entropy was computed as an independent check on the toward-0.5 reading.

Batch checks, software, and use of artificial intelligence

Before any velocity estimate, the distribution of timepoints across BeadChips was cross-tabulated in each cohort to confirm that measurement time is not confounded with sampling time.

Analyses were performed in Python 3.11 and R 4.5.3. Analysis code and derived tables are deposited (see Data Availability). A large language model (Claude, Anthropic) was used to assist in drafting prose and in constructing and reviewing analysis scripts. All analyses were executed by the author, and all numerical results, interpretations, and citations were verified by the author, who takes full responsibility for the content.

Table 1. Characteristics of the two longitudinal cohorts

Characteristic	LSADT (GSE73115)	Health ABC (GSE130748)
Country	Denmark	United States
Array platform	Illumina 450K	Illumina EPIC
Specimen	Whole blood	Buffy coat
Arrays deposited (n)	180	37
Participants paired (n)	86	17 (16 at 60 mo)
Age at intake (y)	73-82	Eighth decade
Interval (mo)	120 (homogeneous)	12-60 (heterogeneous)
Panel coverage	48/48	48/48
Deconvolution reference coverage	713/713	666/713 (93.4%)
Timepoints randomized across chips	Yes (15 chips, 2 plates)	Not assessed
Same-year replicate pairs (n)	8 (distinct chips)	0
Incident disease outcomes	None recorded	Not used

Note. LSADT = Danish Longitudinal Study of Aging Twins; Health ABC = Health, Aging and Body Composition study; 450K = HumanMethylation450 BeadChip; EPIC = MethylationEPIC BeadChip. LSADT participants comprise 43 twin pairs.

Results

Burden increased over 10 years before adjustment

In LSADT, weighted burden rose over the 120-month interval at +0.000518 units/month (SE, 0.000084; t = 6.17; 95% CI, +0.000353 to +0.000682; P = 2.2×10^{-8}). Median burden was 0.686 (range, 0.367-1.074) at baseline and 0.727 (range, 0.416-1.297) at follow-up, a mean increase of 0.062 units over 10 years. The count of loci deviating more than 2 SD from reference rose by a mean of 0.92 loci. Treating each twin pair as one independent unit (effective n = 43) widened the interval to +0.000285 to +0.000751 (P = 1.3×10^{-5}); the result survives that worst-case correction.

In Health ABC the point estimate was concordant in sign and magnitude but was not statistically significant, at +0.000435 units/month in the 60-month arm (SE, 0.000488; t = 0.89; P = .39) and +0.000342 units/month in the full 17-participant model with elapsed time entered per participant (SE, 0.000468; t = 0.73; P = .48). The 95% CI in the 60-month arm spans approximately −0.00052 to +0.00139 and therefore excludes almost nothing.

We note this asymmetry explicitly because it constrains how the two cohorts may be combined. Inverse-variance pooling returns +0.000516 units/month (95% CI, +0.000354 to +0.000678) with Q = 0.028, P = .87, and I^2 = 0%. That apparent absence of heterogeneity is trivially satisfied: LSADT carries approximately 97% of the weight, and an interval as wide as Health ABC's cannot disagree with any plausible value. Health ABC does not independently confirm the LSADT drift estimate; it fails to contradict it, which is a weaker statement, and we treat it as such throughout.

Adjustment for concurrent composition change removes two thirds of the velocity

Adding change in deconvolution-estimated cell fractions to the LSADT model reduced velocity to +0.000167 units/month (SE, 0.000103; t = 1.62; P = .11), retaining 32.2% of the unadjusted magnitude and no longer distinguishable from zero. Dropping eosinophils as the compositional reference to obtain a full-rank design returned an identical intercept (+0.020018) and P = .106, confirming that the rank deficiency of the closed-composition design does not drive the result (Table 2).

In Health ABC the corresponding adjusted estimates were +0.000574 units/month (P = .46; 60-month arm) and +0.000307 units/month (P = .69; full model). Because unadjusted velocity in this cohort was not significantly different from zero, the retained-fraction statistic is a ratio of two near-zero quantities, and the analysis withholds a verdict. Health ABC cannot test the adjustment question.

Table 2. Burden velocity before and after adjustment for concurrent change in leukocyte composition

Cohort and model	Velocity (units/mo)	SE	t	P	R^2
LSADT, unadjusted	+0.000518	0.000084	6.17	2.2×10^{-8}	—
LSADT, unadjusted, twin-corrected	+0.000518	0.000119	4.35	1.3×10^{-5}	—
LSADT, composition-adjusted (min-norm)	+0.000167	0.000103	1.62	.108	0.364
LSADT, composition-adjusted (full rank)	+0.000167	0.000103	1.62	.106	0.364
Health ABC 60-mo, unadjusted	+0.000435	0.000488	0.89	.387	—
Health ABC 60-mo, composition-adjusted	+0.000574	—	—	.459	0.639
Health ABC all pairs, unadjusted	+0.000342	0.000468	0.73	.476	—
Health ABC all pairs, composition-adjusted	+0.000307	—	—	.69	0.580

Note. Velocity is expressed in weighted burden units per month. Retained fraction in LSADT is 32.2%. Standard deviation of Δburden fell from 0.0934 to 0.0745 in LSADT (20.3% reduction) and from 0.1172 to 0.0704 in Health ABC (39.9%). Technical standard deviation of a burden difference, estimated from 8 same-participant replicate pairs on distinct chips, is 0.0328 (95% CI, 0.0217-0.0667), placing 88% of the LSADT variance in Δburden outside measurement error. Health ABC composition-adjusted estimates are reported for completeness; because unadjusted velocity in that cohort is not significantly different from zero, no retained-fraction verdict is issued.

 Worst-case correction treating each twin pair as one independent unit (effective n = 43); 95% CI, +0.000285 to +0.000751.

 Eosinophils dropped as compositional reference. Intercept identical to the minimum-norm fit (+0.020018).

 Coefficient of determination of change in burden regressed on change in estimated leukocyte composition.

Composition change explains a large and replicated share of burden change

Independently of the velocity question, change in cell composition accounted for a substantial fraction of the variance in burden change in both cohorts: $R^2 = 0.364$ in LSADT and $R^2 = 0.639$ in the Health ABC 60-month arm (0.580 across all 17 participants). Adjustment reduced the standard deviation of Δburden from 0.0934 to 0.0745 in LSADT (20.3%) and from 0.1172 to 0.0704 in Health ABC (39.9%).

The direction of compositional change was the expected immunosenescent pattern in both cohorts and on both array platforms. In LSADT, mean fraction changes were CD4+ T −0.0349, neutrophils +0.0357, monocytes +0.0150, B cells −0.0090, CD8+ T −0.0044, natural killer −0.0012, and eosinophils −0.0011. In Health ABC the same lymphoid-to-myeloid direction appeared with larger magnitude: neutrophils +0.0814, CD4+ T −0.0355, natural killer −0.0256, and CD8+ T −0.0198 (Figure 2). Summed across the four lymphoid populations, the loss was −0.0495 in LSADT and −0.0870 in Health ABC, offset almost exactly by myeloid gain (+0.0507 and +0.0879). Identifiable coefficients relative to the eosinophil reference in LSADT were CD4+ T −3.18 (t = −3.44), CD8+ T −3.54 (t = −3.49), natural killer −3.15 (t = −3.10), B cell −2.75 (t = −2.74), neutrophil −2.62 (t = −2.87), and monocyte −1.32 (t = −1.35).

This replication across two array platforms, two specimen preparations, and two national cohorts is the most robust finding in the analysis, and it is worth separating from the velocity result. Whether or not a composition-independent drift exists, composition change is demonstrably a first-order contributor to longitudinal change in a blood-referenced burden score.

Measurement error does not explain the residual

The eight same-participant, same-year replicate pairs available in LSADT, all on physically distinct BeadChips, gave a mean absolute burden difference of 0.0261 and hence a technical standard deviation of a burden difference of 0.0328 (95% CI, 0.0217-0.0667). Against the total observed SD of Δburden in LSADT of 0.0934, this implies a biological SD of 0.0874 and places 88% of the variance in 10-year burden change outside measurement error. Even at the pessimistic upper bound on technical SD, the biological share remains 49%.

This has two consequences. First, the composition-adjusted null is not an artifact of a noisy measurement: participants genuinely differ in how much their burden changes over a decade, which is the necessary precondition for such a measure to carry individual-level information. Second, it sets a floor on individual-level interpretation. Twice the technical SD is 0.0655 units, so at the observed population mean rate an individual would require approximately 127 months for a change to clear measurement error — a constraint on the minimum useful interval between repeat tests that follows directly from the replicate data rather than from assumption.

Change runs toward intermediate methylation rather than along the panel axis

Regressing per-locus mean Δβ on locus centrality returned a negative coefficient in both cohorts (LSADT, −0.01161; SE, 0.00528; t = −2.20; P = .033; model R^2 = 0.340; Health ABC, −0.01866; SE, 0.01421; t = −1.31; P = .20; R^2 = 0.106), indicating that β moved toward 0.5 over the interval (Table 3). Panel Shannon entropy confirmed this independently in LSADT, rising by a mean of 0.00321 (SD, 0.01451; t = 2.05; P = .043; 51 of 86 participants increased). The Health ABC entropy change was null and underpowered (−0.00205; t = −0.68; P = .50; 9 of 17 increased).

The competing predictor — movement along the panel's own hyper/hypo direction vector — was positive but did not survive scrutiny. In LSADT its coefficient was +0.00292 (P = .067), and excluding loci whose reference mean lies outside [0.10, 0.90] and then [0.15, 0.85] weakened it progressively to P = .11 and P = .17. The centrality coefficient moved the other way over the same exclusions, strengthening from −0.01161 to −0.01655 (P = .0059) and −0.01874 (P = .0076), with model R^2 rising from 0.340 to 0.423. A floor artifact would have produced the opposite pattern. The two predictors are correlated (r = −0.642; P = 8.7×10^{-7}) but separable, with a variance inflation factor of 1.70.

The 10-year change is therefore best described as an increase in disorder rather than as coordinated movement along a directional axis.

Table 3. Per-locus regression of mean change in β on locus centrality and panel direction, with boundary sensitivity

Cohort and locus set	Loci (n)	Centrality coefficient	P	Direction coefficient	P	Model R^2
LSADT, all loci	48	−0.01161	.033	+0.00292	.067	0.340
LSADT, reference mean in [0.10, 0.90]	42	−0.01655	.0059	—	.112	0.423
LSADT, reference mean in [0.15, 0.85]	35	−0.01874	.0076	—	.171	0.359
Health ABC, all loci	48	−0.01866	.196	+0.00255	.545	0.106

Note. Centrality is the external reference mean minus 0.5. A negative coefficient indicates movement of β toward 0.5 over the interval, the opposite of the prediction from differential specimen degradation. No dose-response was detected across the 10-year and 5-year storage gaps (z = 0.50; P = .62). The two predictors are correlated (r = −0.642; P = 8.7×10^{-7}) but separable (variance inflation factor, 1.70). Mean Δβ across the 48 loci was +0.00368 in LSADT and −0.00115 in Health ABC. LSADT values average technical replicate arrays within a participant-year. Panel Shannon entropy rose in LSADT (mean, +0.00321; t = 2.05; P = .043) and was null in Health ABC (−0.00205; P = .50).

Differential storage duration and batch structure are excluded

Baseline specimens were stored 10 and 5 years longer than follow-up specimens in LSADT and Health ABC respectively. Progressive degradation predicts a positive centrality coefficient and a dose-response across the two storage gaps. The observed coefficient was negative in both cohorts, and the cohort with the longer gap had the smaller coefficient, with no dose-response detectable (z = 0.50; P = .62). Mean Δβ across the 48 loci was +0.00368 in LSADT and −0.00115 in Health ABC. LSADT values average technical replicate arrays within a participant-year. Differential storage is rejected on both grounds.

In LSADT, the 180 arrays were distributed across 15 BeadChips and 2 plates, every one of which carried both 1997 and 2007 specimens, so sampling time is randomized with respect to measurement batch and the observed change cannot be a chip effect.

Discussion

In two independent cohorts of older adults, most of the apparent decade-scale change in a blood-referenced DNA methylation burden score was attributable to concurrent change in leukocyte composition. After adjustment, the composition-independent remainder in the better-powered cohort was 32% of the unadjusted magnitude and could not be distinguished from zero. Meanwhile, the compositional contribution itself replicated robustly across two array platforms, two specimen preparations, and two countries.

The logic of the within-subject design is what fails here, and it fails for a specific and generalizable reason. Differencing removes confounders that are constant within a participant. Leukocyte composition in later life is not constant; it changes in a consistent lymphoid-to-myeloid direction across participants over exactly the intervals at which longitudinal methylation is measured. A confounder that co-varies with the outcome is not removed by differencing — it is carried into it. Any longitudinal blood methylation biomarker that has not adjusted for concurrent composition change has not established that it is measuring per-cell epigenetic state.

Our finding confirms and quantifies the observation made by Bartlett and colleagues in the Health ABC sample itself (12), who reported extensive change in cell composition over time in participants who went on to develop cancer, including a decline in CD8+ T-cell proportion and a rise in granulocytes — the same direction we measure here. Where that report established the phenomenon, the present analysis measures its share: two thirds of a decade-scale velocity in the larger cohort, with the remainder indistinguishable from zero and separable from measurement error by replicate-derived bounds. Two features make this a stronger result than the P values alone suggest. Estimated cell proportions carry error, and errors-in-variables attenuates their coefficients, biasing the test toward the conclusion that drift survives adjustment. The adjustment nonetheless removed two thirds of the velocity. Separately, the twin correction, which roughly doubles the standard errors, makes the composition-adjusted null more secure, not less.

The directional analysis sharpens the mechanistic reading. The residual change is entropic — β moving toward intermediate values, panel entropy rising — rather than coordinated movement along the panel's directional axis, and the centrality effect strengthens rather than weakens when boundary loci are excluded. This is consistent with the established picture of age-related methylation as an increase in stochastic variability (19-21, 23). It carries a specific design implication: because entropy rises in everyone, an elevated rate of undirected change carries no condition-specific information by itself. A candidate signal must be an excess of directional movement over and above the entropy background, not merely a faster drift rate.

A further observation concerns panel size rather than locus choice. In related work on these same data, a between-study reference offset measured at approximately 23% when computed genome-wide expanded to approximately 140% when computed on this 48-locus panel. Averaging over hundreds of thousands of probes damps technical and batch variation; averaging over 48 does not. A small locked panel is structurally more exposed to study and batch effects than a genome-wide summary, independent of which loci it contains — a consideration for any assay intended to run under a fixed, locked locus set.

Our findings sit alongside a growing literature arguing that composition-aware construction is a prerequisite rather than a refinement for blood methylation measures (10, 13, 14, 26-28). The practical consequences for assay design are concrete: specimen preparation must be locked per participant across timepoints; absolute cross-sectional thresholds calibrated on one preparation should not be transferred to another; and composition adjustment should be reported as a primary analysis rather than a sensitivity check. There is also a plainer implication worth stating. If a large part of what a longitudinal methylation measure detects is a shift in leukocyte proportions, then a complete blood count with differential measures the same construct directly and far more cheaply, and any methylation-based approach must articulate what it adds beyond that.

Limitations

The central limitation is one of power, and it cuts against overinterpretation in both directions. This is a failure to demonstrate composition-independent drift, not a demonstration of its absence: the 95% CI on the adjusted LSADT velocity spans approximately −0.000035 to +0.000369 units/month, so a real composition-independent drift up to roughly two thirds of the unadjusted estimate remains compatible with these data. A larger cohort with three or more draws per participant is required to settle it.

Health ABC contributes corroboration of the compositional finding but cannot test the velocity question, since it showed no significant unadjusted drift to explain. Its 17 pairs give one participant with a 12-month interval extreme leverage, and the pooled velocity estimate should be read as LSADT alone.

Both cohorts comprise older adults, aged 73 to 82 years at intake in LSADT and in the eighth decade in Health ABC. Whether composition dominates longitudinal methylation change at younger ages, where immunosenescent remodeling is slower, is untested here. LSADT participants are twins, so genetic non-independence is present even after the pairwise correction, and the cohort carries no incident disease outcomes.

Methodologically, deconvolution-estimated proportions are estimates rather than measurements; the Health ABC reference carried 666 of 713 probes across a platform change; the two directional predictors are correlated (VIF, 1.70) so their separation, while supported by the boundary-sensitivity analysis, is not definitive; and the healthy reference means underlying the burden score derive in part from a cohort that contributed to the original selection of these loci. That last circularity affects absolute burden values but not the within-participant differences reported here. Finally, this analysis uses one 48-locus panel; the

extent to which the quantitative partition transfers to genome-wide summaries or to other panels is an open question we have not addressed.

Conclusion

Within-participant differencing does not control leukocyte composition, because composition drifts alongside the outcome. In the better-powered of two longitudinal cohorts of older adults, adjusting for concurrent composition change removed two thirds of the apparent decade-scale drift in a blood-referenced methylation burden score and left a remainder indistinguishable from zero, while the compositional contribution replicated across platforms, preparations, and countries. Longitudinal blood methylation biomarkers should report composition-adjusted estimates as a primary result.

Declarations

Funding

This work was supported by Alercell, Inc. No external or public funding was received.

Conflict of Interest

The author is the founder, chief executive officer, and chief scientific officer of Alercell, Inc., a company developing DNA methylation-based diagnostics for hematologic malignancy. The findings reported here constrain the interpretation of blood-based DNA methylation measures generally, including approaches under development at Alercell.

Ethics

This study is a secondary analysis of de-identified, publicly deposited data. No institutional review board approval was required and none was sought. No human participants were recruited and no specimens were collected.

Data Availability

All primary data are publicly available from the Gene Expression Omnibus under accessions GSE73115, GSE130748, and GSE35069 (references 41-43). Analysis code and derived result tables are deposited at Zenodo DOI.10.5281/zenodo.21986630, released under CC BY 4.0 and the MIT License. The complete 48-locus panel specification, including probe identifiers, genome coordinates, directions, reference means, and weights, is provided as eTable 1 in the Supplement.

Acknowledgments

The author thanks the investigators of the Danish Longitudinal Study of Aging Twins, the Health, Aging and Body Composition study, and the contributors of the purified-leukocyte reference series for depositing their data publicly.

References

1. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14(10):R115. doi:10.1186/gb-2013-14-10-r115
2. Hannum G, Guinney J, Zhang L, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell. 2013;49(2):359-367. doi:10.1016/j.molcel.2012.10.016
3. Levine ME, Lu AT, Quach A, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018;10(4):573-591. doi:10.18632/aging.101414
4. Bell CG, Lowe R, Adams PD, et al. DNA methylation aging clocks: challenges and recommendations. Genome Biol. 2019;20(1):249. doi:10.1186/s13059-019-1824-y
5. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13:86. doi:10.1186/1471-2105-13-86
6. Reinius LE, Acevedo N, Joerink M, et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. PLoS One. 2012;7(7):e41361. doi:10.1371/journal.pone.0041361
7. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. Genome Biol. 2014;15(2):R31. doi:10.1186/gb-2014-15-2-r31
8. Franceschi C, Campisi J. Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases. J Gerontol A Biol Sci Med Sci. 2014;69(suppl 1):S4-S9. doi:10.1093/gerona/glu057
9. Ferrucci L, Fabbri E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. Nat Rev Cardiol. 2018;15(9):505-522. doi:10.1038/s41569-018-0064-2
10. Jonkman TH, Dekkers KF, Slieker RC, et al. Functional genomics analysis identifies T and NK cell activation as a driver of epigenetic clock progression. Genome Biol. 2022;23(1):24. doi:10.1186/s13059-021-02585-8
11. Tan Q, Heijmans BT, Hjelmborg JV, Soerensen M, Christensen K, Christiansen L. Epigenetic drift in the aging genome: a ten-year follow-up in an elderly twin cohort. Int J Epidemiol. 2016;45(4):1146-1158. doi:10.1093/ije/dyw132
12. Bartlett AH, Liang JW, Sandoval-Sierra JV, et al. Longitudinal study of leukocyte DNA methylation and biomarkers for cancer risk in older adults. Biomark Res. 2019;7:10. doi:10.1186/s40364-019-0161-3
13. Teschendorff AE, Zheng SC. Cell-type deconvolution in epigenome-wide association studies: a review and recommendations. Epigenomics. 2017;9(5):757-768. doi:10.2217/epi-2016-0153
14. Zheng SC, Breeze CE, Beck S, Teschendorff AE. Identification of differentially methylated cell types in epigenome-wide association studies. Nat Methods. 2018;15(12):1059-1066. doi:10.1038/s41592-018-0213-x
15. Salas LA, Koestler DC, Butler RA, et al. An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina HumanMethylationEPIC BeadArray. Genome

Biol. 2018;19(1):64. doi:10.1186/s13059-018-1448-7

16. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363-1369. doi:10.1093/bioinformatics/btu049
17. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. Nucleic Acids Res. 2017;45(4):e22. doi:10.1093/nar/gkw967
18. Chen YA, Lemire M, Choufani S, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics. 2013;8(2):203-209. doi:10.4161/epi.23470
19. Teschendorff AE, Menon U, Gentry-Maharaj A, et al. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. Genome Res. 2010;20(4):440-446. doi:10.1101/gr.103606.109
20. Issa JP. Aging and epigenetic drift: a vicious cycle. J Clin Invest. 2014;124(1):24-29. doi:10.1172/JCI69735
21. Slieker RC, van Iterson M, Luijk R, et al. Age-related accrual of methylomic variability is linked to fundamental ageing mechanisms. Genome Biol. 2016;17(1):191. doi:10.1186/s13059-016-1053-6
22. Hannon E, Knox O, Sugden K, et al. Characterizing genetic and environmental influences on variable DNA methylation using monozygotic and dizygotic twins. PLoS Genet. 2018;14(8):e1007544. doi:10.1371/journal.pgen.1007544
23. Jenkinson G, Pujadas E, Goutsias J, Feinberg AP. Potential energy landscapes identify the information-theoretic nature of the epigenome. Nat Genet. 2017;49(5):719-729. doi:10.1038/ng.3811
24. Marioni RE, Shah S, McRae AF, et al. DNA methylation age of blood predicts all-cause mortality in later life. Genome Biol. 2015;16:25. doi:10.1186/s13059-015-0584-6
25. Chen BH, Marioni RE, Colicino E, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. Aging (Albany NY). 2016;8(9):1844-1865. doi:10.18632/aging.101020
26. Higgins-Chen AT, Thrush KL, Wang Y, et al. A computational solution for bolstering reliability of epigenetic clocks: implications for clinical trials and longitudinal tracking. Nat Aging. 2022;2(7):644-661. doi:10.1038/s43587-022-00248-2
27. Sugden K, Hannon EJ, Arseneault L, et al. Patterns of reliability: assessing the reproducibility and integrity of DNA methylation measurement. Patterns (N Y). 2020;1(2):100014. doi:10.1016/j.patter.2020.100014
28. Tomusiak A, Floro A, Tiwari R, et al. Development of an epigenetic clock resistant to changes in immune cell composition. Commun Biol. 2024;7(1):934. doi:10.1038/s42003-024-06609-4
29. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med. 2014;371(26):2488-2498. doi:10.1056/NEJMoa1408617

30. Genovese G, Kähler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med. 2014;371(26):2477-2487. doi:10.1056/NEJMoa1409405
31. Steensma DP, Bejar R, Jaiswal S, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood. 2015;126(1):9-16. doi:10.1182/blood-2015-03-631747
32. Klutstein M, Nejman D, Greenfield R, Cedar H. DNA methylation in cancer and aging. Cancer Res. 2016;76(12):3446-3450. doi:10.1158/0008-5472.CAN-15-3278
33. Lehne B, Drong AW, Loh M, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. Genome Biol. 2015;16:37. doi:10.1186/s13059-015-0600-x
34. Weidner CI, Lin Q, Koch CM, et al. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. Genome Biol. 2014;15(2):R24. doi:10.1186/gb-2014-15-2-r24
35. Pal S, Tyler JK. Epigenetics and aging. Sci Adv. 2016;2(7):e1600584. doi:10.1126/sciadv.1600584
36. Teschendorff AE, Marabita F, Lechner M, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2013;29(2):189-196. doi:10.1093/bioinformatics/bts680
37. Scheer F. Compartment mismatch produces near-perfect apparent discrimination in blood-referenced DNA-methylation classifiers: a public-cohort reanalysis of hematologic-malignancy series. Research Square. Preprint posted online July 27, 2026. doi:10.21203/rs.3.rs-10472780/v1
38. Scheer F. Epigenetic age estimates in blood shift by decades with cell population: a within-donor variance partition across two platforms and two study designs. Research Square. Preprint posted online July 29, 2026. doi:10.21203/rs.3.rs-10503047/v1
39. Scheer F. 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. Preprint posted online July 23, 2026. doi:10.21203/rs.3.rs-10397076/v1
40. [dataset] Tan Q, Christiansen L, et al. Epigenetic drift in the aging genome: a ten-year follow-up in an elderly twin cohort. Gene Expression Omnibus; 2016. Accession GSE73115.
41. [dataset] Bartlett AH, Liang JW, Sandoval-Sierra JV, et al. Longitudinal study of leukocyte DNA methylation and biomarkers for cancer risk in older adults. Gene Expression Omnibus; 2019. Accession GSE130748.
42. [dataset] Reinius LE, Acevedo N, Joerink M, et al. Differential DNA methylation in purified human blood cells. Gene Expression Omnibus; 2012. Accession GSE35069.

Figures

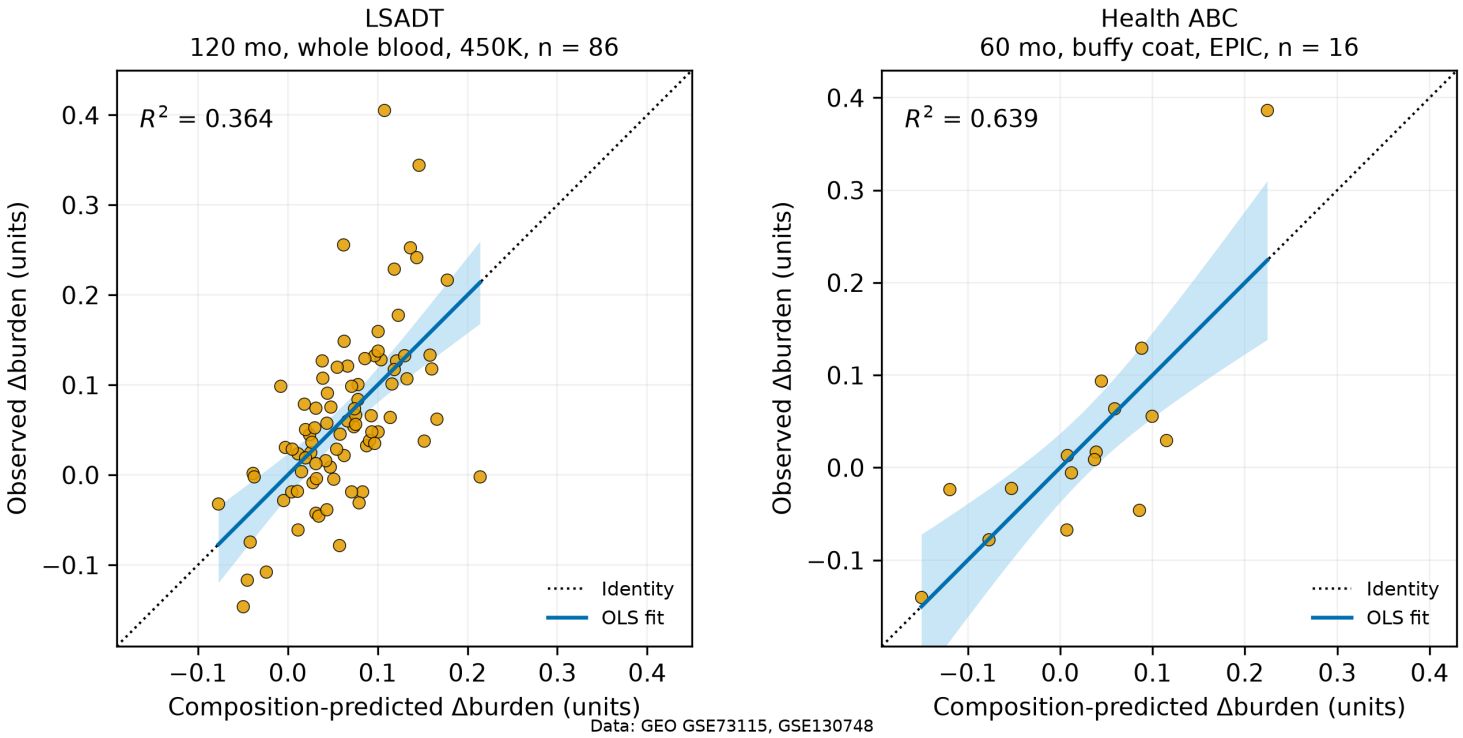

Figure 1

Observed change in weighted burden score plotted against the change predicted from concurrent leukocyte composition alone, for each participant, shown separately for LSADT (120-month interval, whole blood, Illumina 450K, n = 86) and Health ABC (60-month interval, buffy coat, Illumina EPIC, n = 16). Predicted values are fitted from a linear model of Δburden on the six non-reference Δ cell fractions, eosinophils dropped as the compositional reference. Dotted lines mark identity; solid lines show the ordinary least-squares fit with 95% confidence bands. Coefficients of determination are 0.364 and 0.639 respectively.

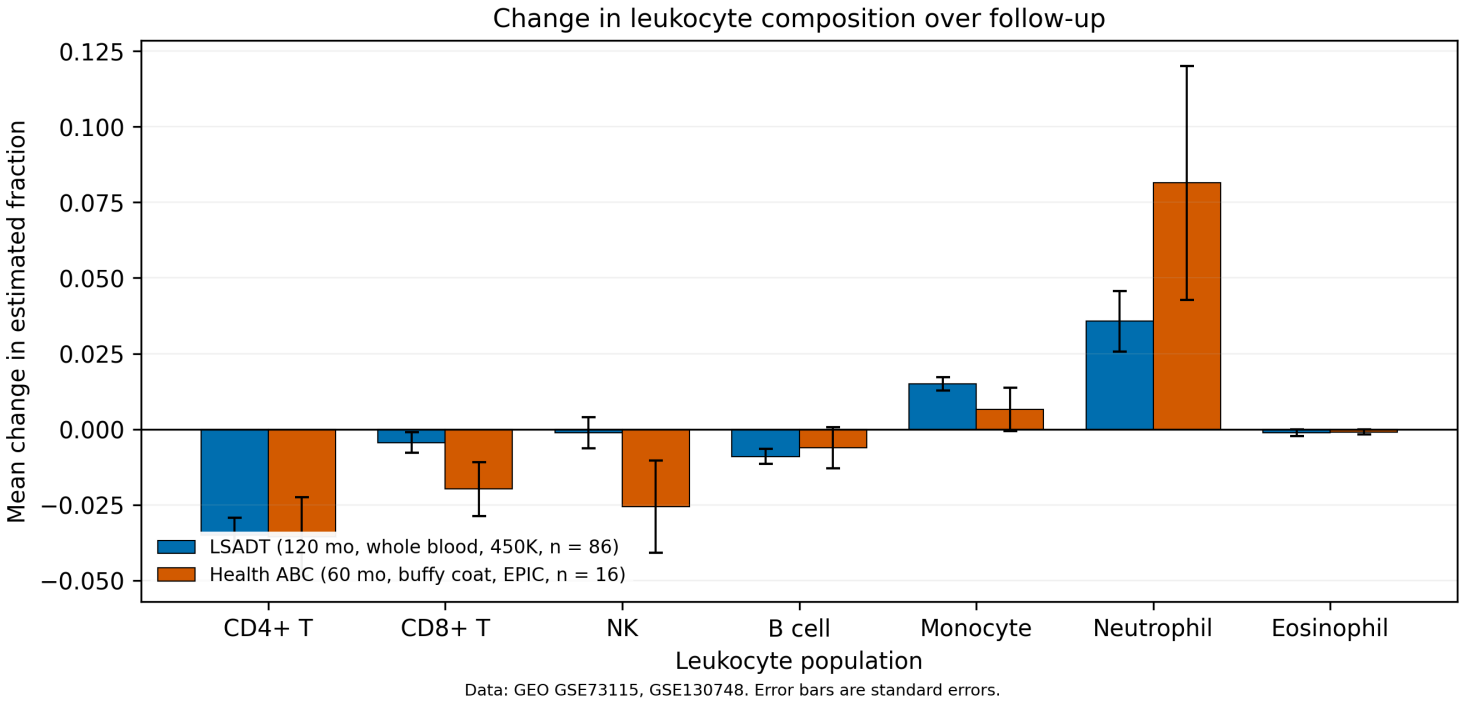

Figure 2

Mean change in estimated fraction of each leukocyte population over the follow-up interval in both cohorts. Bars show mean change; error bars show standard errors. Both cohorts show the same lymphoid-to-myeloid direction — loss of CD4+ T, CD8+ T, and natural killer cells with gain of neutrophils — despite differing array platform, specimen preparation, interval length, and country.

Supplementary Files

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

- SupplementaryMaterial.docx