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.

Blast percentage confounds transcriptome to drug-response associations in Beat AML: DHODH and sorafenib as a worked example

Cumhur Gokhan Ekmekci

cgekmekci2@yahoo.com

Independent Researcher https://orcid.org/0000-0002-0967-172X

Research Article

Keywords: leukemia, myeloid, acute, gene expression profiling, antineoplastic agents, drug screening assays, antitumor, confounding factors, epidemiologic, dihydroorotate dehydrogenase, sorafenib, Beat AML, blast percentage

Posted Date: November 1st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11501028/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

Beat AML pairs ex vivo dose response for 125 compounds with transcriptomes from the same patients and is widely mined for markers of drug sensitivity. Those analyses adjust for general drug sensitivity, differentiation state, proliferation and genotype, but not for the blast content of the specimen, and a recent four-cohort study concluded that such an adjustment was not warranted.

Across 12,000 gene-by-drug pairs, 35 per cent of the associations significant under the usual adjustment set lost significance when blast percentage was added, against 3.7 per cent gained. The pairs came from 800 genome-wide genes and the 15 most tested compounds, with the sample set fixed per drug so that nothing turns on a change in n. Blast percentage correlated with both transcript abundance and residual drug response. The comparable figures for tumor purity in solid tumors are about 14 and 11 per cent.

A worked example shows the cost to a candidate marker. Under the usual model DHODH expression was the strongest correlate of sorafenib response among the genes tested, sorafenib ranked first of 125 drugs (rho − 0.259, FDR 6.5×10^{-7}) and the empirical p against an expression-matched genome-wide null was below 0.001. With blast percentage added, rho fell to − 0.186 (n 414), sorafenib to rank four and the empirical p to 0.012. What survived was selective: five of the six other FLT3-directed compounds showed nothing.

Blast percentage belongs in the adjustment set for these analyses. It is already in the diagnostic record and costs nothing to include.

Introduction

When a transcript is tested against a drug response in Beat AML, which covariates belong in the model? The answer has settled into a short list by convention rather than by argument, and the list decides which genes are reported as markers.

Beat AML provides ex vivo dose response for 125 compounds alongside exome and transcriptome data from the same patients, with clinical annotation [1, 2]. It is the uncommon resource where a transcriptional correlate of drug response can be tested at patient level rather than inferred from cell lines, and it is used that way often, from single-gene association scans to machine-learning models of drug response.

Two confounders are handled almost routinely in that literature. Some samples are more easily killed by everything, which is dealt with by working from residual rather than raw response. And drug response in AML is governed broadly by differentiation state, which the Beat AML investigators themselves showed and which deconvolution work has since made quantitative [2, 3]. A third is not handled. Samples differ in the proportion of blasts they contain, and the non-blast fraction dilutes both the transcriptome and the ex vivo readout [4, 5]. A recent four-cohort consistency study examined this directly and reported no

association between marrow blast percentage and response for the drugs it tested, concluding that including it as a covariate was not warranted [8].

Composition itself is not ignored here. Deconvolution of Beat AML transcriptomes against single-cell references has been used to reconstruct differentiation stage, estimate leukemic stem cell fractions, predict ex vivo resistance and compare marrow with blood specimens [3, 16, 17]. What has not followed is its use as a covariate when a single gene is tested against a drug, which is the analysis that produces candidate markers.

The problem is not new in oncology. In more than 10,000 solid tumors, purity was shown to confound co-expression, clustering and differential expression, and adjusting for it made about 14 per cent of differentially expressed genes lose significance while about 11 per cent gained it [6]. Estimators of purity from expression followed [7], including deconvolution against a healthy marrow reference in AML [3]. The caution has not been carried into ex vivo drug-response analysis, where the non-blast fraction affects both the transcriptome and the assay readout.

This matters because the candidate markers that come out of such scans are what later get proposed for validation, and a covariate left out of the model does not announce itself in the output. A marker inflated by composition looks exactly like a marker.

The analysis reported here began elsewhere, as a search for transcriptional markers of sorafenib sensitivity, and produced a candidate strong enough to be worth defending: DHODH, the mitochondrial enzyme of de novo pyrimidine synthesis, better known in AML as the target of differentiation-inducing inhibitors [9, 10]. Testing that candidate against the obvious objections is what surfaced the blast problem, and the size of the problem turned out to be the more useful result.

Methods

Data

Beat AML waves 1 to 4 were used throughout: normalised expression, probit dose-response curve fits, exome mutation calls and the harmonised clinical table, all from the public release [2]. Drug response is area under the curve, where lower values indicate greater sensitivity. Samples with both expression and drug data numbered 520, from 479 patients; analysis was restricted to the 125 compounds tested in at least 100 of them. Marrow blast percentage was available for 398 of the 520 and a specimen-matched blast value for 437. TCGA-LAML expression and clinical data came from the PanCancer Atlas release through cBioPortal [11], with RSEM values log2-transformed.

Models

Each drug's AUC was regressed on the sample's median AUC across all other compounds and the residual used. A differentiation axis was built as the difference between a monocytic score (LYZ, CD14, CD68, FCN1, VCAN, MAFB, CSF1R, CEBPB) and a primitive score (CD34, KIT, SPINK2, MSI2, PROM1,

DNMT3B), each the mean of z-scored expression, with a proliferation score from MKI67, PCNA, TOP2A, CCNB1 and AURKA. These and FLT3-ITD status form the base model. The extended model adds blast percentage, in two versions. Marrow blast is used for the scale comparison and the genome-wide calibration, since it covers the widest set of specimens and keeps those analyses on one footing. A specimen-matched version, marrow blast for marrow and blood blast for blood specimens, is used for the primary DHODH estimate and brings in 38 more specimens. The text says which applies where. Within FLT3-ITD strata, ITD status drops out and the other covariates are retained. Covariates were partialled out on ranks by least squares.

The scale comparison

To ask whether the problem generalises, 800 protein-coding genes were sampled genome-wide by reservoir sampling and tested against the 15 most widely measured compounds, giving 12,000 pairs. For each drug the sample set was fixed to the specimens both models could use, so that a difference between models could not arise from a difference in n. Significance was taken at Benjamini-Hochberg FDR 0.1 computed across the whole set of pairs within each model.

Calibration

Choosing the strongest of many genes inflates significance, so the DHODH association was calibrated against a null built the same way, and separately under each model. Protein-coding genes were sampled genome-wide and those within 0.75 log2 units of DHODH mean expression were run against sorafenib under the identical adjusted model. The empirical p is the fraction of matched genes reaching a correlation at least as negative as the observed one. Running the calibration under one model while reporting an estimate from another would answer a different question, so both are given.

Statistics

Associations are Spearman correlations on ranks, intervals are Fisher z, and multiplicity within a scan is handled by Benjamini-Hochberg. Survival used Cox proportional hazards with the same covariates, as hazard ratio per standard deviation of expression. Remission compared patients recorded as achieving complete response with those recorded as refractory by Mann-Whitney, excluding unknown responses rather than counting them as failures. Analyses were run in Python with pandas, scipy and statsmodels.

Results

Blast percentage is associated with both sides of the analysis

A covariate matters when it correlates with the predictor and with the outcome. Marrow blast percentage does both. It was recorded for 398 specimens and correlated with DHODH expression at rho +0.353 (p 4.2×10^{-13}). Among the 376 of those that also have a sorafenib measurement, it correlated with residual sorafenib response at rho −0.285 (p 1.8×10^{-8}). Peripheral blood blast percentage, available for 314 samples, behaved the same way (+0.366 and −0.225). Both relationships are shown in Figure 1A and 1B.

The direction is what one would expect mechanically rather than biologically: a specimen with more blasts carries more of the proliferating compartment, so proliferation-linked transcripts read higher, and the ex vivo assay has more target cells to kill.

It is not a restatement of what the model already contains, correlating with the differentiation axis at rho −0.272 and with proliferation at −0.193.

How many associations change

A confounder that moves one association is a curiosity; one that moves many is a problem with the resource. We therefore repeated the comparison at scale, taking 800 protein-coding genes sampled genome-wide and the 15 most widely tested compounds, and scoring all 12,000 pairs under both models with the sample set fixed per drug.

Under the base model 1,187 of the 12,000 pairs passed FDR 0.1. Under the extended model 801 did. Of the base hits, 416 lost significance, which is 35 per cent, and 30 associations appeared that had not been significant before, which is 3.7 per cent of the extended hits. For comparison, the pan-cancer purity analysis of solid tumors reported about 14 per cent lost and about 11 per cent gained [6]. The effect in Beat AML is larger and much less symmetric.

Two numbers in Figure 2 describe different things. The absolute correlation shrank in 57 per cent of all pairs, which is the behaviour of the whole cloud; 35 per cent is the share of significant associations that crossed back below the threshold, higher because those sat near it to begin with.

It is also not one compound. Dasatinib fell from 192 significant genes to 114, nilotinib from 175 to 113, erlotinib from 52 to 21 and pazopanib from 52 to 15 (Table 2). Sorafenib, 121 to 42, is among the most affected but not an outlier, while trametinib barely moved, 238 to 222, so the correction is not a blanket deflation.

Table 2. Significant gene-drug associations before and after adding blast percentage, by compound. 800 genome-wide genes per drug, FDR 0.1 computed across all 12,000 pairs, sample set fixed per drug.

Drug	Significant, base	Significant, extended	Median |rho| base	Median |rho| extended
Trametinib	238	222	0.086	0.087
Dasatinib	192	114	0.078	0.064
Nilotinib	175	113	0.074	0.066
Sorafenib	121	42	0.065	0.049
Bortezomib	84	63	0.057	0.059
Vandetanib	66	58	0.049	0.049
Rapamycin	55	43	0.052	0.050
Erlotinib	52	21	0.054	0.047
Pazopanib	52	15	0.053	0.045
Crizotinib	34	24	0.045	0.044
Gefitinib	33	20	0.049	0.048
Lapatinib	30	26	0.044	0.045
Axitinib	24	18	0.044	0.044
Sunitinib	24	19	0.043	0.044
Imatinib	7	3	0.034	0.030
All 15 drugs	1,187	801	0.053	0.050

What this does to one candidate marker

Under the base model, DHODH expression against sorafenib response gave rho −0.259 (p 5.2×10^{-9}, n 494). Testing DHODH against the whole panel placed sorafenib first of 125 drugs, with 17 compounds past FDR 0.1 (Figure 3A, Table 1), and the empirical p against 164 expression-matched genome-wide genes was below 0.001 (Supplementary Figure S1A). On those numbers the association would be reported as strong and specific.

Adding blast percentage gives rho −0.186 (p 1.4×10^{-4}, n 414) with the specimen-matched measurement, and −0.155 (p 0.0026, n 376) using marrow blast alone. We take the matched version as primary, because marrow blast percentage is not the right measurement for a blood specimen and the matched variable brings 38 more specimens into the estimate. The panel-wide scan and the calibration were run with marrow blast, so that the estimate and its null are built the same way: there sorafenib falls to rank four of 125, the number of compounds past FDR 0.1 falls from 17 to 7, and the empirical p rises to 0.012 (Figure 3B, Supplementary Figure S1B). Figure 3 and Table 1 report that version throughout. The

specimen-matched value of −0.186 is the one to quote for the size of the association, and we have not recomputed the ranking under it.

Table 1. DHODH against sorafenib under the two models, with the calibration recomputed under each. The extended column uses marrow blast percentage, the measurement for which the genome-wide null was also computed; the specimen-matched estimate is given in the text.

Quantity	Base model	Extended model (+ blast)
rho, sorafenib	−0.259	−0.155
p	5.2×10^{-9}	0.0026
n	494	376
rank among 125 drugs	1	4
FDR	6.5×10^{-7}	0.083
drugs past FDR 0.1	17	7
empirical p, matched genome-wide null	< 0.001	0.012
null 2.5th percentile	−0.210	−0.140
survival, HR per SD	0.847 (p 0.010)	0.821 (p 0.015)

What survives the adjustment

The attenuated association is still selective for sorafenib among FLT3-directed compounds. After blast adjustment, quizartinib gave rho +0.002, gilteritinib +0.103, crenolanib +0.075, sunitinib −0.057 and ponatinib −0.064, none significant. The association held separately in FLT3-ITD-positive patients (rho −0.208, p 0.048, n 91) and ITD-negative patients (−0.134, p 0.024, n 285), with quizartinib and gilteritinib null in both (Supplementary Table S1). Whatever is left is not FLT3 dependence.

Two other explanations were ruled out before blast percentage. Mitochondrial content was the obvious candidate, since DHODH sits on the inner mitochondrial membrane and venetoclax preferentially kills OXPHOS-dependent AML; adding an OXPHOS score, a mitochondrial mass score or both moved the sorafenib estimate from −0.259 to −0.257 (Supplementary Figure S3). The two drug responses correlate at rho +0.249, but conditioning each on the other strengthened them rather than weakening them, to −0.269 and −0.235.

One result runs the other way and should not be smoothed over. Midostaurin, flat in the base model at +0.061, became +0.150 (p 0.005) once blast percentage was included. The scale analysis puts that in proportion: 30 associations were created by the adjustment against 416 removed, so creation is real but uncommon, and the same asymmetry appears in the solid-tumor purity work. We report midostaurin as a caution about the extended model rather than as a finding.

Other objections, and which ones mattered

Four further sources of confounding were tested, each as its own model. Beat AML holds more than one specimen for 34 patients, up to four in one case; restricting to one sample per patient, chosen by a rule fixed before the analysis, gave rho −0.246. Working from raw rather than residual AUC gave −0.212. Adding specimen type gave −0.259 and a relapse indicator gave −0.260. None of these moved the estimate appreciably (Figure 1C, Supplementary Table S2). Blast percentage was the only adjustment that did.

Pathway context

DHODH does not act alone. Under the base model, CAD, UMPS, CTPS1, CTPS2 and RRM1 tracked sorafenib response in the same direction and more weakly, while TYMS and RRM2 did not (Supplementary Table S3). A score from the three de novo steps placed sorafenib first of 125 drugs, venetoclax second and cytarabine fourth; a salvage-pathway score placed the same three at ranks 100, 103 and 83 (Supplementary Figure S2). Dropping DHODH weakened the score and dropping UMPS strengthened it. These analyses were run under the base model and carry the same inflation, so they are context rather than evidence.

Outcome, and a cohort where it does not replicate

In Beat AML, higher DHODH expression accompanied longer overall survival under both models (HR 0.847 per standard deviation, p 0.010 in the base model; 0.821, p 0.015 with blast percentage included) and was higher in patients achieving complete remission than in those recorded as refractory (p 0.001).

This did not replicate. In TCGA-LAML, 149 patients with usable survival and 114 events gave HR 0.913 unadjusted (p 0.385) and 0.927 with differentiation and proliferation included (p 0.524). The direction matches, the effect does not reach significance, and at this sample size the analysis has roughly even odds of detecting an effect of the size seen in Beat AML, so this is a failure to confirm rather than a contradiction. Two things further weaken it as a test: the age field is empty for every patient in that release, and TCGA patients received standard induction rather than sorafenib. Blast percentage could not be modelled there either. No prognostic claim is made.

An analysis that failed

This work began as a ferroptosis hypothesis: that defence architecture would predict response to two drugs hitting different arms of it, sorafenib through system xc− and lovastatin through the mevalonate route to CoQ10. Neither predicted cell held, at rho −0.024 (p 0.59) and +0.100 (p 0.25) respectively (Supplementary Figure S4). Evidence against sorafenib as a system xc− inhibitor [12] makes that unsurprising in retrospect. It is reported because the DHODH candidate came out of that scan, and the number of hypotheses tested around a result is part of how it should be read.

Discussion

The result with the widest reach here is the simplest. Blast percentage correlates with transcript abundance and with ex vivo response in Beat AML, and leaving it out of the adjustment set costs about a third of the associations that would otherwise be reported. That figure comes from 12,000 gene-drug pairs with the sample set held fixed, so it is not an artefact of one gene, one drug or a change in sample size. It is an average rather than a rule: the loss ranged from almost nothing for trametinib to two thirds for sorafenib and pazopanib, so which associations survive depends on the compound.

The kind of problem is not new [6]; the setting and the size are. Ex vivo drug response gives composition a second route to act, because the assay counts the death of whatever cells are present, and published Beat AML analyses do not generally adjust for it. The loss here is two and a half times the solid-tumor figure and far more one-sided, which fits that extra route.

This contradicts the conclusion of a four-cohort consistency study, which found no such association and advised against including the covariate [8]. We cannot reconcile the two from the published description. That study pooled 993 samples from four cohorts in a linear mixed framework and worked mainly from a normalised drug sensitivity score rather than area under the curve, asking which experimental factors shift response between laboratories. That is a different estimand from a within-cohort partial correlation on ranks after removing general drug sensitivity and differentiation state, and a between-laboratory analysis can be insensitive to a covariate that matters within one. It is also possible that the association we see is specific to Beat AML. Either way the disagreement is worth settling, and the cost of including the covariate is small compared with the cost of an inflated marker.

The closest published analysis to our worked example illustrates the problem. Hogeling and colleagues correlated the proteome of primary AML samples with ex vivo sensitivity to a DHODH inhibitor and found cholesterol and lipid programs raised in the sensitive samples, with proliferation showing a trend that did not reach significance [18]. Same class of analysis, same disease, same readout, and composition not modelled there either. We cannot say what their result would look like with a blast term in it; the point is that nobody can. The same applies to our own pathway analyses, which is why they are labelled base-model throughout.

On the worked example, something does survive. After adjustment the DHODH association is weak, rho −0.186 with the specimen-matched blast variable and −0.155 with marrow blast alone, but still selective: five of the six other FLT3-directed compounds show nothing, the sixth moves the other way, and the effect is present in both FLT3-ITD strata, so it is not FLT3 dependence. DHODH is the mitochondrial step of de novo pyrimidine synthesis, the salvage pathway shows no comparable signal, and venetoclax, which suppresses the same step by depolarising the membrane DHODH requires [13,14], behaves as a separate association rather than the same one twice. A mechanism is not established, and nothing here could establish one. Sorafenib has been shown to suppress MCL1 through STAT3 [15], and MCL1 is required for the integrity of the outer mitochondrial membrane, so a route through mitochondrial priming is a hypothesis worth testing rather than an explanation we can offer.

Several explanations for the residual DHODH association compete and can be ranked only by plausibility. It may still be partly a composition measure, since morphological blast percentage is coarse and a gene tracking the proliferating fraction keeps some signal after adjustment for a noisy version of it. It may be pharmacological: de novo pyrimidine synthesis competes with cytarabine, and blocking DHODH restores cytarabine sensitivity in AML models, which fits cytarabine ranking fourth for the de novo score. It may be the MCL1 priming route described above. None is mutually exclusive and none is tested here.

One model accommodates the whole set of results. Specimen composition acts twice, on the transcriptome and on an assay that counts the death of whatever cells are present, so any transcript tracking the blast fraction acquires an apparent association with any drug that kills blasts. Adjusting for composition removes that shared component and leaves whatever is specific to a drug and a gene. What is left in this dataset is small, concentrated in compounds with a mitochondrial or nucleotide connection, and absent from the FLT3 inhibitors, which is what a residual metabolic axis would look like after the composition term has been taken out.

DHODH should not be proposed as a clinical selector on this basis. The adjusted effect explains a few per cent of the variance in an ex vivo surrogate, the prognostic half did not replicate, and expression is not activity: DHODH is regulated post-translationally and bulk mRNA is a loose proxy for flux. The useful version is that de novo pyrimidine capacity is worth measuring directly in a sorafenib-treated cohort, which is an experiment rather than a biomarker.

Our own analysis has a practical fix. Blast percentage is missing for 122 of 520 specimens and is a morphological estimate where it exists, whereas expression-based deconvolution reconstructs AML composition well enough to predict ex vivo resistance [3] and purity estimators perform comparably in solid tumors [7]. A deconvolved blast fraction would be available for every sample and would not depend on an observer.

Limitations

Everything here is correlational. Beat AML is an observational resource, the associations are partial correlations between transcript abundance and an ex vivo assay readout, and no experiment in this work manipulates blast content, DHODH or pyrimidine synthesis. The attenuation we report is therefore a statement about how these data behave under two models, not a demonstration that blast content causes the associations to appear; and the DHODH result is a candidate that would need functional work to become anything more.

Blast percentage is measured by morphology, with observer variation, and is missing for 122 of 520 specimens. That missingness is not random. It falls mainly on peripheral blood specimens, 101 of 228 of which lack a marrow blast value against 15 of 277 marrow aspirates, and the specimens without a value have lower proliferation scores ($p\ 6.0 \times 10^{-8}$) and lower DHODH expression ($p\ 0.005$). Complete-case analysis therefore rests on an assumption that is demonstrably false here. The specimen-matched

variable recovers 39 of those samples and is used as primary for that reason, but it does not solve the problem.

Ex vivo AUC is a surrogate for clinical response. Technical factors including time to testing, culture medium and viability assay produce systematic differences between laboratories [8], and no clinical sorafenib cohort was available here.

DHODH was selected as the best gene in a scan, so the rank-one result for sorafenib is partly a restatement of how it was found. The expression-matched null is the test of that and is reported under both models. The midostaurin result also shows that adjustment can create associations as well as remove them, as the purity work found.

The scale analysis used 800 genes and 15 compounds rather than the whole transcriptome and the whole panel. The proportions should be read as estimates with sampling error, not as exact figures.

The pathway and ferroptosis analyses were run under the base model and carry the same inflation. Five hypotheses were tested and discarded in the work that produced this one, and the calibration is meant to be read with that in mind.

Conclusions

What was not known before this analysis is how much the omission of blast percentage costs in Beat AML. It costs about a third of the associations that pass FDR 0.1, with the sample set held fixed, and the loss is one-sided: 416 associations removed against 30 created. The covariate is not a restatement of differentiation or proliferation, and no other objection we tested moved the estimates.

Three predictions follow, each testable without new patients. Published Beat AML marker claims recomputed with blast percentage should lose significance at roughly the rate reported here, and more for compounds in the dasatinib to pazopanib range than for trametinib. A blast fraction deconvolved from expression should reproduce the attenuation on the 398 specimens with a morphological value and extend it to the other 122; if it does not, the morphological measure is carrying something other than composition. And if the residual DHODH association is pharmacological, then in a sorafenib-treated cohort pre-treatment de novo pyrimidine capacity measured directly should separate responders from non-responders, with no such effect for quizartinib or gilteritinib in the same patients.

Declarations

The author declares no competing interests. No funding was received. This study used only publicly available, de-identified data and required no ethical approval.

Data and code availability

All data are public. Beat AML waves 1 to 4 are distributed by the study investigators [2] and TCGA-LAML through cBioPortal [11]. Analysis scripts and derived tables are deposited at Zenodo, doi:10.5281/zenodo.23644010.

References

1. Tyner JW, Tognon CE, Bottomly D et al (2018) Functional genomic landscape of acute myeloid leukaemia. Nature 562:526–531. 10.1038/s41586-018-0623-z
2. Bottomly D, Long N, Schultz AR et al (2022) Integrative analysis of drug response and clinical outcome in acute myeloid leukemia. Cancer Cell 40:850–864. 10.1016/j.ccell.2022.07.002
3. Karakaslar EO, Severens JF, Sánchez-López E et al (2024) A transcriptomic based deconvolution framework for assessing differentiation stages and drug responses of AML. NPJ Precis Oncol 8:105. 10.1038/s41698-024-00596-9
4. Kaspers GJ, Veerman AJ, Pieters R et al (1994) Mononuclear cells contaminating acute lymphoblastic leukaemic samples tested for cellular drug resistance using the methyl-thiazol-tetrazolium assay. Br J Cancer 70:1047–1052. 10.1038/bjc.1994.446
5. Cucchi DGJ, Groen RWJ, Janssen JJWM, Cloos J (2020) Ex vivo cultures and drug testing of primary acute myeloid leukemia samples: current techniques and implications for experimental design and outcome. Drug Resist Updat 53:100730. 10.1016/j.drup.2020.100730
6. Aran D, Sirota M, Butte AJ (2015) Systematic pan-cancer analysis of tumour purity. Nat Commun 6:8971. 10.1038/ncomms9971
7. Revkov E, Kulshrestha T, Sung KW-K, Skanderup AJ (2023) PUREE: accurate pan-cancer tumor purity estimation from gene expression data. Commun Biol 6:394. 10.1038/s42003-023-04764-8
8. Willoch-Allen K, Zhao Z, Ayuda-Durán P et al (2026) A multi-center study on the consistency of drug sensitivity testing in patients with acute myeloid leukemia. NPJ Precis Oncol 10:294. 10.1038/s41698-026-01480-4
9. Sykes DB, Kfoury YS, Mercier FE et al (2016) Inhibition of dihydroorotate dehydrogenase overcomes differentiation blockade in acute myeloid leukemia. Cell 167:171–186. 10.1016/j.cell.2016.08.057
10. Christian S, Merz C, Evans L et al (2019) The novel dihydroorotate dehydrogenase (DHODH) inhibitor BAY 2402234 triggers differentiation and is effective in the treatment of myeloid malignancies. Leukemia 33:2403–2415. 10.1038/s41375-019-0461-5
11. Hoadley KA, Yau C, Hinoue T et al (2018) Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. Cell 173:291–304. 10.1016/j.cell.2018.03.022
12. Zheng J, Sato M, Mishima E, Sato H, Proneth B, Conrad M (2021) Sorafenib fails to trigger ferroptosis across a wide range of cancer cell lines. Cell Death Dis 12:698. 10.1038/s41419-021-03998-w

13. Gu X, Tohme R, Goldfinger M et al (2020) Venetoclax inhibition of pyrimidine synthesis guides methods for integration with decitabine or 5-azacytidine that are non-myelosuppressive. Blood 136(Suppl 1):26–27. 10.1182/blood-2020-143200
14. Mishra R, Zokaei Nikoo M, Veeraballi S, Singh A (2024) Venetoclax and hypomethylating agent combination in myeloid malignancies: mechanisms of synergy and challenges of resistance. Int J Mol Sci 25:484. 10.3390/ijms25010484
15. Wang R, Xia L, Gabrilove J, Waxman S, Jing Y (2016) Sorafenib inhibition of Mcl-1 accelerates ATRA-induced apoptosis in differentiation-responsive AML cells. Clin Cancer Res 22:1211–1221. 10.1158/1078-0432.CCR-15-0663
16. Caliskan G, Pawitan Y, Vu TN (2024) Similarities and differences of bone marrow and peripheral blood samples from acute myeloid leukemia patients in terms of cellular heterogeneity and ex-vivo drug sensitivity. eJHaem 5:721–727. 10.1002/jha2.961
17. Coudray A, Forey R, Bejar Haro B et al (2026) Computational deconvolution of drug sensitivity via single-cell-to-bulk transcriptome mapping. PLoS Comput Biol 22:e1014862. 10.1371/journal.pcbi.1014862
18. Hogeling SM, Sternadt D, La Rose N et al (2026) Potent synergy of DHODH and SREBP inhibition in acute myeloid leukemia via disruption of cholesterol and lipid metabolism. Haematologica 111:2319–2329. 10.3324/haematol.2025.287918

Figures

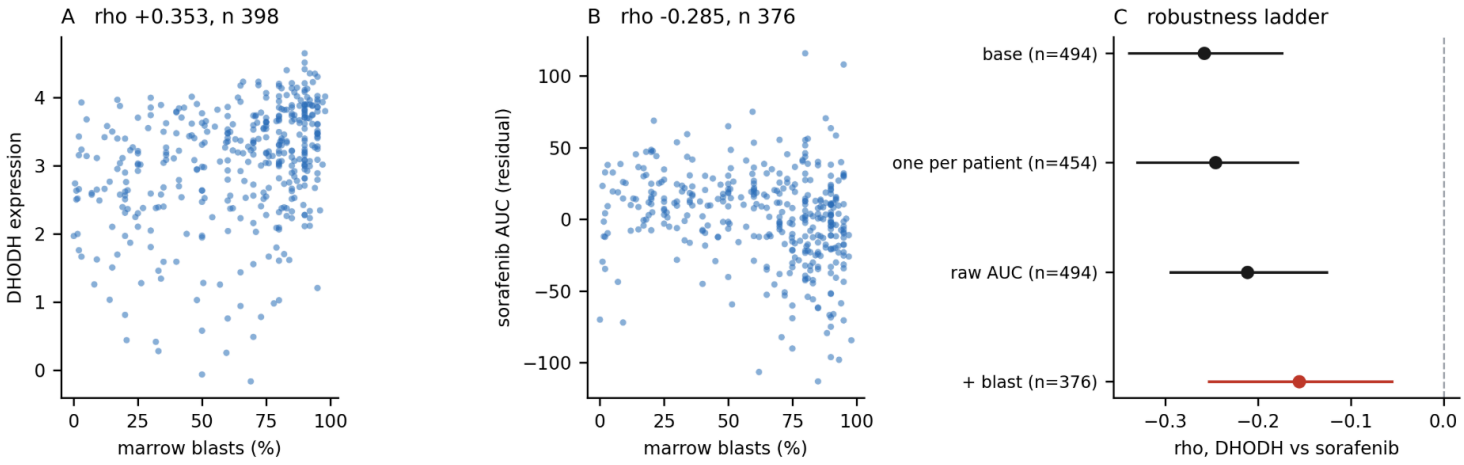


Figure 1

Blast percentage as a confounder. (A) Marrow blast percentage against DHODH expression, rho +0.353, n 398. (B) Marrow blast percentage against residual sorafenib response, rho −0.285, n 376. A covariate that correlates with both sides of an analysis will bias it, and this one does. (C) The DHODH to sorafenib association under four models with 95% Fisher z intervals. Restricting to one specimen per patient, working from raw rather than residual AUC, and adding specimen type or a relapse indicator leave the

estimate where it was; adding blast percentage (red) halves it. The full robustness series is in Supplementary Table S2.

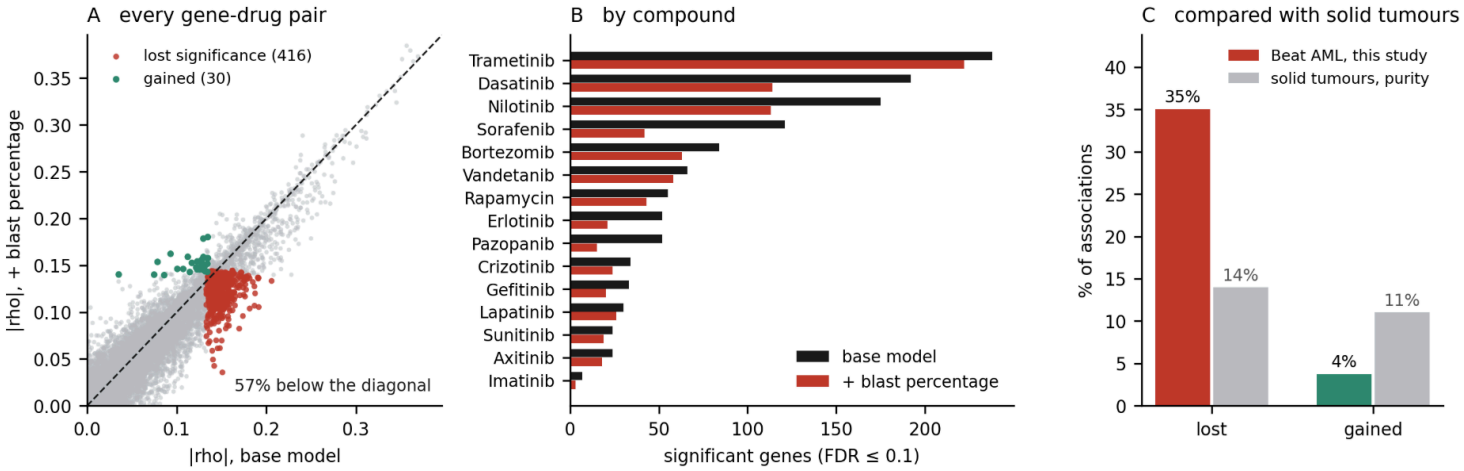


Figure 2

Attenuation at scale. (A) Absolute adjusted rho under the base model against the same quantity under the extended model, for all 12,000 gene-drug pairs. Points below the diagonal shrank once blast percentage entered the model, which is 57 per cent of them. Red, the 416 associations significant under the base model that stopped being significant; green, the 30 that became significant. (B) Significant associations per compound before and after. (C) The proportion lost and gained here, beside the figures reported for tumor purity in solid tumors [6].

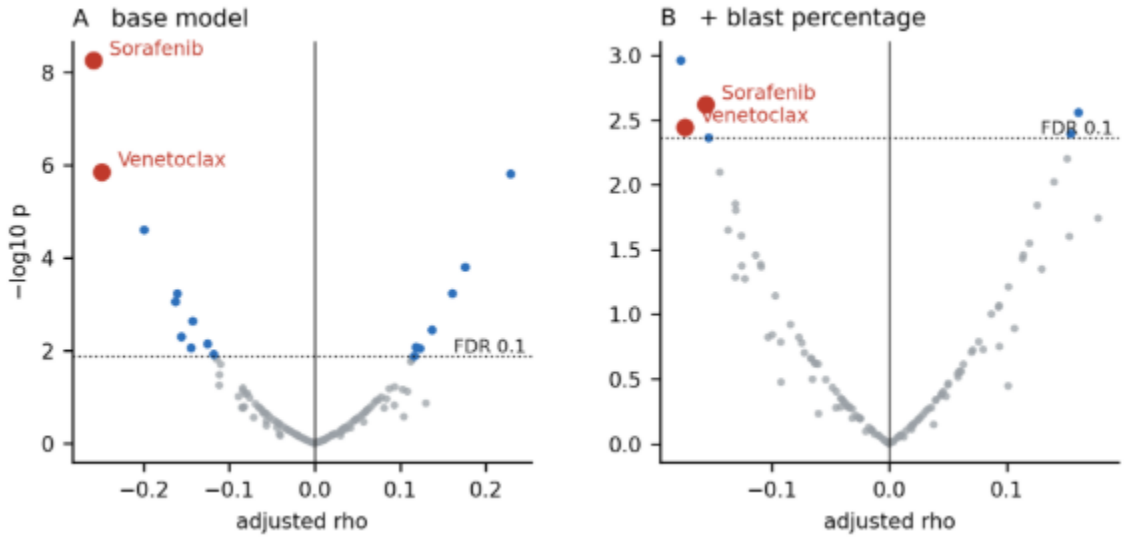


Figure 3

One marker under both models. Adjusted Spearman rho against −log10 p for DHODH tested against 125 compounds. Grey, not significant; blue, FDR ≤ 0.1; red, the two labelled compounds; dotted line, FDR 0.1. (A) Base model: sorafenib first, 17 compounds past threshold, the scale running to eight orders of

magnitude. (B) With blast percentage added: sorafenib fourth, seven compounds past threshold, the scale running to three.

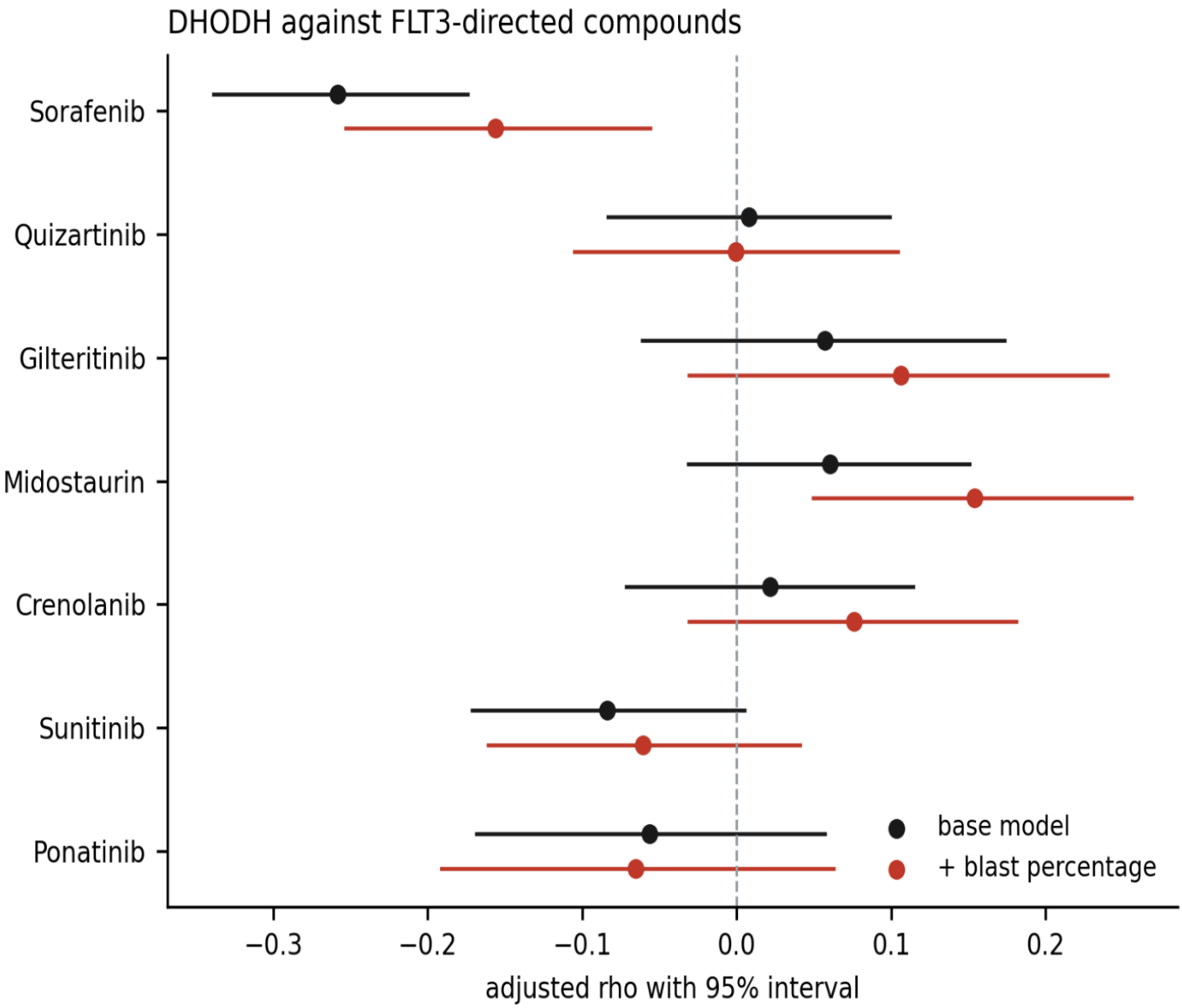


Figure 4

DHODH against FLT3-directed compounds under both models, with 95% Fisher z intervals. Black, base model; red, with blast percentage added. Sorafenib is the only compound whose interval excludes zero under both. Midostaurin crosses from null to positive once blast percentage is included.

Supplementary Files

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

- SupplementaryBlastConfoundingEkmekci2026v6.docx