Clinical Exposure Normalization
Restores Translational Predictiveness of In Vitro Drug Response

Eugene F Douglass Jr, [1,*]; Chinmay Joshi, [1] Tobin Paez,[1] Elizabeth A. Hughes,[1] Emily Hannan,

[1] Department of Pharmaceutical and Biomedical Sciences, University of Georgia, 250 W. Green Street, Athens, GA 30602

* Correspondence: eugene.douglass@uga.edu

ABSTRACT

Traditional laboratory methods quantify drug-efficacy based on dose-response metrics (e.g. IC50 or AUC), whereas clinical oncology defines drug-efficacy based on time-response metrics (e.g. progressive or stable disease). These measures often diverge, limiting translation between bench and bedside. We developed NCI1970-Meta, a harmonized dataset of 49,002 evaluable patients covering 30 drugs and 18 cancers, to provide the first large-scale benchmark linking *in vitro* potency, clinical pharmacokinetics, and Phase 2–3 outcomes data.

Raw *in vitro* metrics values showed poor correlation with clinical response, systematically underestimating antimetabolites and overvaluing outliers such as gemcitabine. By contrast, normalizing potency to clinically achievable exposure (e.g. the maximum plasma concentration (Cmax)) restored concordance with trial outcomes. Biomarker analyses, which compare patients treated with the same drug, were largely robust to this dichotomy: 234 curated biomarkers showed concordance across clinical and *in vitro* datasets. In addition, curated transporter knockout data validated a parallel flux model of additive resistance, explaining the limited success of multi-drug resistance inhibitors for drug-transporter enzymes.

These findings establish exposure-normalized potency as a simple, actionable framework for translational oncology. We recommend screening to the unbound Cmax and normalizing potency metrics by exposure to improve the fidelity of functional assays, reduce bias in AI models, and refine biomarker discovery. Beyond cytotoxic agents, NCI1970-Meta provides a general framework for aligning laboratory pharmacology with clinical outcomes.

INTRODUCTION:

Clinical-Laboratory "efficacy" dichotomy. High-throughput drug screening has generated millions of *in vitro* dose–response curves across hundreds of cancer cell lines and FDA-approved agents.[1-4] These assays typically define activity by the concentration that reduces cell viability by 50% (IC50), or area under the curve (AUC) creating a potency-based view of drug action (Fig. 1A). In contrast, clinical oncology defines efficacy by tumor shrinkage over time, measured as overall response rate (ORR) during Phase 2–3 trials (Fig. 1B).[5] The disconnect between laboratory dose–response metrics and clinical time–response outcomes has limited the translational utility of preclinical assays (Fig. 1C).[6]

This gap is reflected in the arc of drug development. Dose-response studies guide dose-escalation studies in phase 1 human studies where the therapeutic range (TR) is determined, bounded by a minimum needed for efficacy (TR_{min}) and the maximum plasma concentration (C_{max}) achievable in patients (Fig 1A).[7,8] Yet, it is only in Phase 2–3 trials that true therapeutic benefit is assessed through tumor regression (Fig. 1B-C). Linking these stages remains challenging because clinical single-agent ORR data are fragmented, often buried in decades-old monographs or meta-reviews, and rarely digitized for systematic analysis. [9] Without this clinical anchor, potency values measured *in vitro* can easily mislead translational research: drugs that reach high plasma levels may appear inactive in laboratory screens capped at 10 μM,[1-4] while agents that cannot achieve therapeutic exposures may look artificially potent *in vitro*. [10]

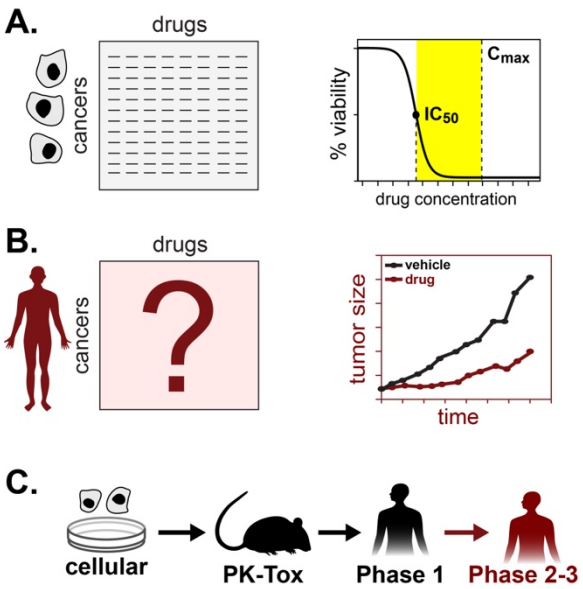

Figure 1. Conceptual overview data-gap (Clinical response rates) filled by this study. (A) Most *in vitro* drug response data is defined by *dose–response metrics* such as AUC or IC50, (B) In contrast, clinical response data is defined by *time-response metrics* which is currently not available in a digitized form. (C) Drug development proceeds from cellular assays to animal PK–toxicity, Phase 1 dose-escalation, and finally Phase 2–3 efficacy trials. We align these stages by comparing *in vitro* IC50 values with, Phase 1 drug-exposure and clinical response rates.

***Laboratory* Single-Agent Data**. Efforts to standardize laboratory pharmacology began with the NCI-60 screen in the early 1990s, which profiled thousands of compounds annually across 60 genomically characterized cancer cell lines.[1] This initiative evolved into larger resources such as the Cancer Therapeutics Response Portal (CTRP),[2] the Genomics of Drug Sensitivity in Cancer (GDSC),[3] and the PRISM barcoded-pool platform, which now collectively cover more than 1,000 cell lines and over 1,000 drugs.[4] These dose-response centric resources (Fig 1A) represent the backbone of preclinical pharmacogenomics, enabling systematic mapping of genotype–drug associations and fueling machine-learning models of sensitivity.[3,11-13] Yet despite their breadth, they remain unanchored to Phase 2–3 efficacy data, which continues to define regulatory approval and clinical adoption.[5] In practice, this disconnect has made it difficult to evaluate whether *in vitro* potency metrics such as IC50 or AUC are truly predictive of real-world therapeutic benefit.

***Clinical* Single-Agent Data**. A parallel limitation arises on the clinical side. Most modern oncology trials investigate drug combinations, leaving relatively few contemporary studies of single-agent efficacy.[14] The critical foundation for many chemotherapy regimens was established in the 1960s–1980s, during an era of national single-agent trials conducted under the auspices of the National Cancer Institute (NCI). [9] These studies enrolled tens of thousands of patients and directly measured monotherapy response rates, but their results remained scattered across monographs, internal reports, and review articles. Although these data shaped the field's understanding of drug activity and underpinned combination regimens still in use today, they were never systematically digitized or reconciled with laboratory measures.[9] As a result, a key translational resource has remained inaccessible to modern computational and biomarker-based oncology (Fig 1B).

To address this gap, we created *NCI1970-Meta*, a harmonized dataset of 49,002 patients across 786 single-agent chemotherapy trials, covering 30 drugs and 18 cancer types. By digitizing and reconciling these historical data – including the landmark 1970 NCI monograph *Single Agents in Cancer Chemotherapy* and subsequent meta-reviews[9] – we provide the first large-scale, quantitative map connecting *in vitro* potency, Phase 1 pharmacokinetic boundaries, and Phase 2-3 efficacy outcomes. This dataset allows us to systematically benchmark the relationship between laboratory potency metrics and clinical response, clarifying when and why translational failures occur.

Using this resource, we demonstrate that *exposure-normalized potency (i.e. IC_{50}/C_{max} or AUC/C_{max})* restores predictive alignment between laboratory and clinical efficacy. This simple adjustment corrects for systematic drug-class biases, explains anomalies such as the overestimation of gemcitabine, and offers actionable design principles for both experimental pharmacology and computational oncology. In particular, we show that two immediate changes can improve translational fidelity: (i) screening drugs up to their clinically achievable C_{max} to avoid false negatives, and (ii) normalizing IC_{50} values by C_{max} or the therapeutic minimum (TR_{min}) to enable fair cross-drug comparisons.

Beyond benchmarking potency metrics, NCI1970-Meta enables systematic testing of chemotherapy biomarkers and resistance models. Our analyses confirm that drug-metabolism and efflux pathways consistently outperform direct drug targets in predicting response to cytotoxic chemotherapies, and we present a mechanistic model that explains how redundant transporter networks blunt the efficacy of multidrug resistance (MDR) inhibitors. Together, these applications illustrate how historical clinical data can inform modern challenges, from AI model training to organoid-based functional precision medicine.

In summary, this study establishes the first comprehensive lab-to-clinic bridge for cytotoxic chemotherapy. By integrating legacy clinical trials with pharmacokinetic and *in vitro* screening data, we provide both a historical foundation and a forward-looking framework to align laboratory assays with real-world therapeutic outcomes (Fig 1).

RESULTS:

The NCI1970-meta Dataset.

Our study began with the discovery of the 1970 NCI monograph Single Agents in Cancer Chemotherapy, which compiled results from the institute's national monotherapy trials spanning 1955–1970.[9] Although many of these studies were never formally published, the monograph preserved responder counts and evaluable patients in a consistent format. We digitized this resource and organized it into a matrix of drug–cancer pairs, yielding 19,958 evaluable patients across 786 trials covering 30 drugs and 18 tumor types. However, this initial dataset was highly sparse, with 58% of possible drug–cancer combinations untested (Fig. S1).

To reduce this sparsity, we incorporated data from 48 meta-reviews published between 1970 and 2000. These reviews aggregated Phase II–III single-agent trials across specific drugs or cancer types and often provided larger, more statistically robust cohorts than the original NCI data. To avoid duplication, we applied a rule-based algorithm that selected the most comprehensive dataset available for each drug–cancer pair (Fig. S2).

The final reconciled dataset—termed NCI1970-Meta—includes 49,002 evaluable patients treated with 30 chemotherapies across 18 malignancies. The dataset records both total patients and responders for each trial, enabling standardized calculation of overall response rate (ORR) across a theoretical matrix of 540 combinations.[5,15,16] While common cancers such as breast, lung, colorectal, and ovarian were well represented, many rarer cancers remained sparsely tested. [17] This curated resource provides the first harmonized framework for single-agent chemotherapy response, serving as a bridge between historical clinical outcomes and modern laboratory drug screens(Fig. S3).

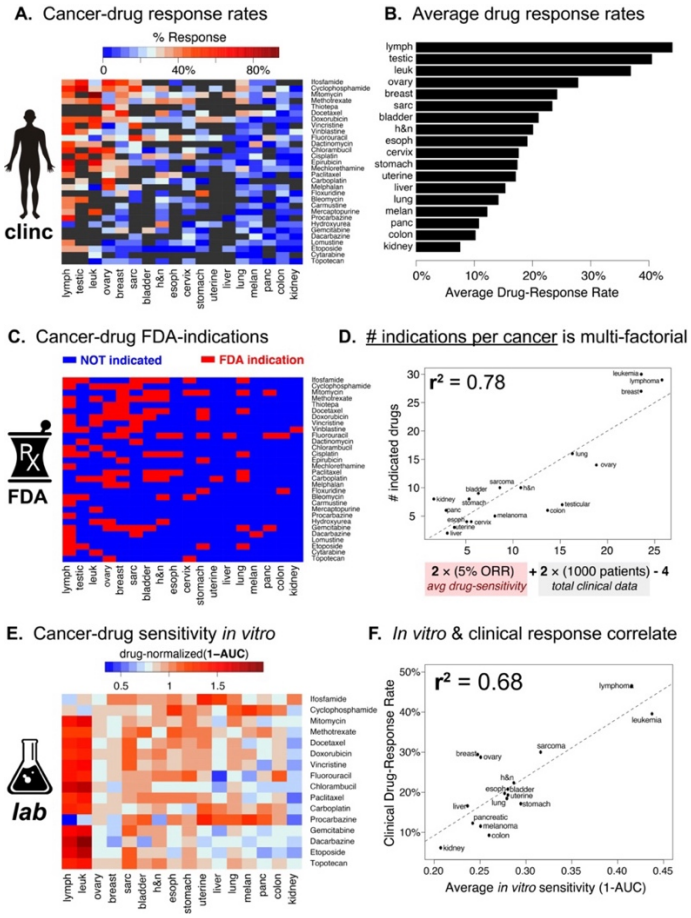

Figure 2. "Rosetta stone" analysis: using NCI1970-meta to compare FDA-indication and laboratory drug-sensitivity data (A). Heatmap of ORR data with drugs and cancers sorted by highest average ORR **(B)** Average ORR across cancer-types **(C)** FDA-indications[14] visually mirror NCI1970-meta ORR data **(D)** The number of indications per cancer appears to be a function of cancer chemosensitivity (in red) and the total amount of clinical data (in grey). **(E).** *In vitro* drug-sensitivity appears to show similar lineage dependence on **(F)** *In vitro* drug-sensitivity accounts for 68% of the variance in clinical drug-sensitivity.

Clinical vs Laboratory Potency Metrics.

Clinical response mirrors FDA approvals. We next compared clinical response rates from NCI1970-Meta with FDA-approved indications and large-scale *in vitro* screening datasets to evaluate the translational relevance of laboratory potency metrics (Fig. 2). Across 30 drugs and 18 cancer types, overall response rates (ORRs) revealed a consistent clinical activity profile. Hematologic malignancies exhibited the highest ORRs, while gastrointestinal cancers—particularly pancreas, colon, and kidney—were the least responsive (Fig. 2A–B). This pattern closely resembled the current landscape of FDA single-agent indications (Fig. 2C).[14] Indeed, the number of indications per cancer type correlated with both the observed clinical response rate and the total number of evaluable patients ($R^2 = 0.78$), suggesting that approval reflected both efficacy and trial feasibility (Fig. 2D).

Raw *in vitro* potency is systematically biased. We then compared these clinical profiles with preclinical sensitivity data from the CTRP and GDSC screens, which report area-under-the-curve (AUC) metrics across hundreds of cancer cell lines.[2,3] Raw AUC values revealed strong drug-class biases: natural products such as taxanes and anthracyclines appeared broadly potent, whereas antimetabolites and electrophiles were systematically underestimated (Fig. S4A–B). These biases obscured lineage-specific differences seen in clinical ORRs and FDA approvals.

Normalization restores concordance. When *in vitro* sensitivity was normalized within each drug to remove baseline potency effects, tissue-specific activity patterns emerged that more closely aligned with clinical

ORRs (Fig. 2E, S4C-D). At the lineage level, normalized *in vitro* sensitivity explained 68% of the variance in clinical response (R^2 = 0.68, Fig. 2F, S5). This analysis highlights the importance of adjusting laboratory potency metrics to account for intrinsic drug-class differences before comparing activity across tumor types. Together, these results show that clinical efficacy patterns are faithfully captured in NCI1970-Meta and mirrored in FDA approvals,[14] but not in raw laboratory potency scores.[1-4] Correcting for drug-specific biases substantially improves alignment, setting the stage for deeper analyses that incorporate pharmacokinetic exposure (e.g. C_{max}).

Drug-Specific Biases in Laboratory Potency Metrics.

While normalizing *in vitro* data within each drug improved alignment with clinical outcomes (Fig. 2), systematic discrepancies remained at the level of individual compounds. To clarify these drug-specific biases, we directly compared clinical **overall response rates** (ORRs) from NCI1970-Meta with area-under-the-curve (AUC) values from large-scale laboratory screens[1-4] and published therapeutic range data (Fig. 3A).[10,18]

Raw AUC values fail to predict clinical efficacy. Across 30 drugs, raw *in vitro* potency metrics showed no global correlation with clinical ORRs (Fig. 3B).[1-4] Instead, drugs clustered into three distinct groups (Fig. 3C). First, natural products such as anthracyclines and taxanes displayed intermediate AUCs but aligned well with their clinical activity. Second, antimetabolites and electrophiles appeared inactive *in vitro*, despite broad clinical utility. Third, gemcitabine emerged as a striking outlier: highly potent in laboratory assays yet associated with only modest ORRs in patients. These systematic biases reveal why raw AUC values, even when averaged across cancer types, cannot serve as reliable predictors of clinical benefit (Fig 3B-C).

Clinical pharmacokinetics explain many discrepancies. We hypothesized that these distortions reflected differences in clinically achievable drug exposures. Indeed, many antimetabolites and electrophiles routinely reach unbound plasma concentrations ($C_{max,u}$) of 100–400 μM in patients—far exceeding the 10 μM ceiling used in CTRP, GDSC, or PRISM assays.[1-4] For example, the clinical $C_{max,u}$ of 5-fluorouracil is ~426 μM, ifosfamide ~431 μM, and carboplatin ~135 μM, each well above the laboratory dose range.[10] As a result, their apparent inactivity *in vitro* may be an artifact of assay design, not a true lack of efficacy. Comparisons of laboratory AUC's with clinical $C_{max,u}$ values partially explained these distortions, showing highly significant correlation relative to ORR values (Fig. 3D).

Therapeutic range metrics resolve outliers. Gemcitabine remained an exception as it featured the lowest AUC's and among the highest $C_{max,u}$ values (Fig. 3D). Recognizing that $C_{max,u}$ often represents the ceiling of the therapeutic range, we examined the correlation between *in vitro* AUC values and a recently published compilation of therapeutic minima plasma concentrations (TR_{min}) (Fig. 3E).[18] Strikingly the therapeutic minimum appeared to almost perfectly correlate with laboratory derived AUC values (Fig 3E).

Together, these analyses demonstrate that *in vitro* dose-response metrics like AUC, largely reflect therapeutic exposure (C_{max} or TR_{min}) not clinical efficacy (ORR) and therefore care should be taken as a surrogate for clinical response. This is a critical insight as these dose-response metrics are often used as surrogates for clinical efficacy.[3,11-13]

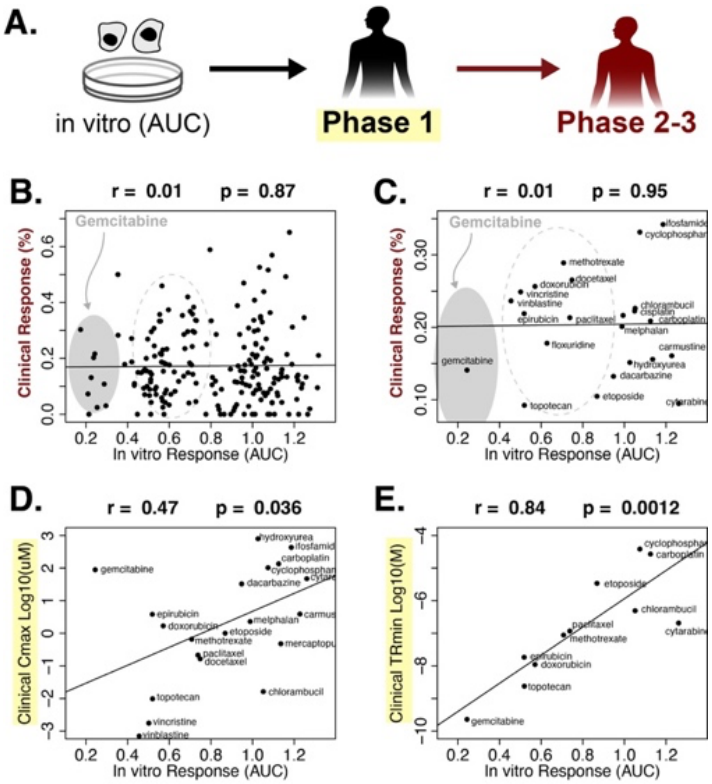

Figure 3. **Drug-specific biases in laboratory potency metrics and their resolution by clinical exposure. (A)** Translational framework: *in vitro* dose–response (AUC/IC50) informs discovery and Phase 1 PK (Cmax), while Phase 2–3 trials define efficacy by tumor regression (ORR). **(B)** Raw *in vitro* AUC values show no correlation with clinical response rates (ORR). **(C)** Averaged across cancer types, drug classes separate: natural products align with clinical efficacy, antimetabolites/electrophiles are underestimated, and gemcitabine is an outlier. **(D)** Clinical maximum plasma concentrations (Cmax) partially explain drug-specific biases.[10] **(E)** Incorporating therapeutic minimum concentrations (TRmin) fully restores concordance, resolving the gemcitabine anomaly and aligning *in vitro* potency with clinical efficacy.[18]

In addition, several practical laboratory practice recommendations arise from this analysis. First, Limiting dose–response curves to sub-therapeutic ranges systematically underestimates classes such as antimetabolites (Fig S4), while ignoring therapeutic range metrics exaggerates the promise of outliers like gemcitabine (Fig 3B-C). Second, normalizing potency values to clinical exposure – whether C_{max} or TR_{min} – partially restores translational relevance and provides a rigorous framework for interpreting *in vitro* data in the context of real-world dosing.

Applications of the Framework: Biomarker Evaluation

Based on our framework, comparisons of biomarkers for a ***single-drug*** [19-21] are expected to be more robust to the laboratory-clinical divide than multi-drug prioritization studies[6,22-25] since exposure ($TR_{min}/C_{max,u}$) is constant across patients. A secondary goal of NCI1970-Meta was to test whether laboratory-derived biomarkers of cytotoxic chemotherapy (CC) sensitivity translate clinically. Hundreds of candidate biomarkers have been reported for CC over five decades – covering drug targets, metabolizing enzymes, transporters, and downstream correlates[26] —but few have been clinically validated due to the scarcity of digitized response datasets.[27,28]

Comprehensive Biomarker Assessment. As illustrated in Figure 4A, this effort represents a direct comparison between 50 years of laboratory literature focused on CC biomarkers and 50 years of clinical trial data compiled in NCI1970-Meta. We used RNA expression as the unifying modality, based on extensive evidence that transcriptomic signatures outperform DNA mutations and other -Omic readouts in predicting drug sensitivity.[3,11] Biomarker annotations were compiled from DrugBank and extended via targeted literature searches (see Methods), and were categorized into four mechanistic classes: ***direct drug-binding targets***, **drug-*metabolizing enzymes***, ***drug transporters***, and ***indirect downstream targets*** (Fig 4E, S6).[29,30] Their predictive power was then tested across two transcriptomic resources: The Cancer Genome Atlas (TCGA) tumors (clinical domain) and the Cancer-Cell Encyclopedia (CCLE) cell lines (*in vitro* domain).[31-33]

Composite biomarker sets were consistently predictive. Across nearly all drugs evaluated, we observed consistent positive correlations between biomarker expression and response in both clinical and *in vitro* settings—validating the robustness of these literature-derived features (Fig 4C-D). Notably, composite scores that integrated all biomarkers outperformed individual markers for every drug. To dissect the relative contribution of each mechanistic class, we color-coded the overall predictive strength of each drug's biomarker set according to the dominant contributing category. This approach highlighted both shared and drug-specific patterns of biomarker utility and provided a rigorous framework for quantifying their translational consistency across platforms (Fig 4C-D).

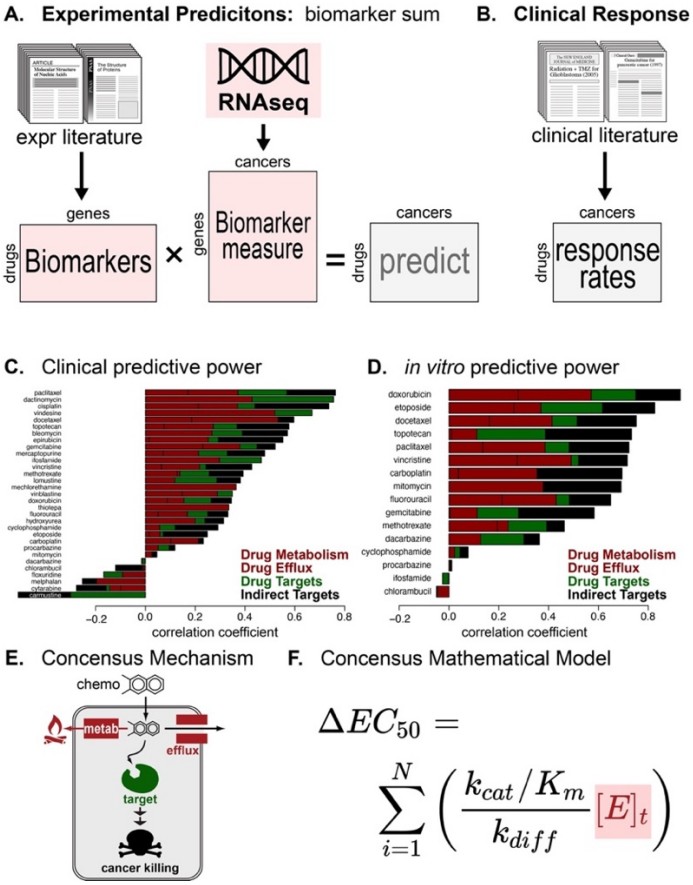

Figure 4. **Laboratory biomarkers sets effectively predict *in vitro* and clinical drug-response (A)** Experimental CC-efficacy predictions are made by simply summing the mRNA expression of all compiled literature biomarkers **(B)** Clinical response rate data was compiled from the clinical literature as described above **(C)** Biomarkers sums across TCGA correlate with clinical response rates data across drugs from Figure 1A **(D)** Biomarkers sums across CCLE correlate with *in vitro* sensitivity from Figure 1E **(E)** Consensus mechanism for CC's compiled from the experiment literature as described above **(F)** Consensus mathematical model for the effect of drug efflux and drug metabolism enzymes on drug potency as defined in Equation 2.

Drug-metabolism and efflux enzymes consistently exhibited stronger predictive power than classical ***drug targets*** across our dataset, suggesting that non-target determinants play a large role in shaping clinical and laboratory response. To further explore this finding, we focused on the three drug classes with the highest clinical and *in vitro* predictiveness—taxanes, vinca alkaloids, and anthracyclines. Using the procedure outlined in Figure S7, we evaluated the correlation of individual biomarkers within these classes and found strong

agreement between *in vitro* and in vivo predictive power (Figure S8). Importantly, metabolism and efflux genes frequently outperformed direct targets in explaining drug sensitivity variance. This reinforces the view that chemotherapy response reflects a multi-gene phenotype, rather than simple target expression. [19,28,34] To quantify the independent contributions of these biomarkers, we applied multi-linear regression and identified top features with statistically significant associations to drug response (Figure S9), consistent with previous studies showing strong performance of linear models in modeling chemotherapy sensitivity.[3,11,12]

A Comprehensive Chemoresistance Model. Given the strong predictive power of linear regression models of drug-metabolism and efflux we next sought to derive a mathematical framework that explains their influence on chemotherapy efficacy. Starting from first principles, we extended classical pharmacodynamic theory by incorporating Michaelis–Menten-type kinetics to model the effect of efflux and metabolic enzyme expression on drug potency.[35-37] As described in the Supporting Information (Fig S10), we solved for IC_{50} as a function of enzyme abundance $[E]_t$ and the ratio of enzymatic to membrane diffusion rate constants:[38]

$$\Delta IC_{50} = \sum_{i=1}^{N} \left(\frac{k_{cat}/K_m}{k_{diff}} \cdot [E]_t \right) \qquad (1)$$

We validated the model in two ways. First, the predicted IC_{50} shifts agreed closely with observed changes in drug potency following genetic induction or inhibition of efflux pathways *in vitro* (Fig S11).[41] Second, we tested the model against an independent dataset comprising blood–brain barrier (BBB) permeability measurements for 76 clinically approved drugs in mice.[40] In this dataset, brain-to-plasma concentration ratios were reported for wild-type, Abcb1a-/-, Abcg2-/-, and double-knockout strains.[40] According to our model, the barrier imposed by two transporters should be additive: the effect of the double knockout should approximate the sum of the individual knockouts. Strikingly, this prediction held across the entire drug panel, with summed single-knockout shifts matching the wild-type to double-knockout differences almost exactly (Fig 5).

Single vs redundant transporters. This framework helps reconcile early enthusiasm for multidrug resistance (MDR) inhibitors with their later clinical failure.[42] In one influential study, deletion of the efflux pump Abcb1a in mice caused ivermectin toxicity at doses more than 20-fold lower than wild type (Fig. 5A).[39] This dramatic phenotype reflects the fact that ivermectin is almost exclusively a substrate of Abcb1a; thus, eliminating a single transporter had catastrophic effects. By contrast, our curated dataset of 76 FDA-approved drugs tested in Abcb1a-/-, Abcg2-/-, and double-knockout mice showed that most compounds are substrates of both transporters (Fig. 5B). Thus, elimination of a 10x fold change enzyme will only result in a 2x change in BBB-ratio if transporter redundancy is present (Fig 5B). Most cases corresponded to these conditions where single-knockouts had limited effects, relative to double knockouts yielded shifts approximating the sum of both.[40]

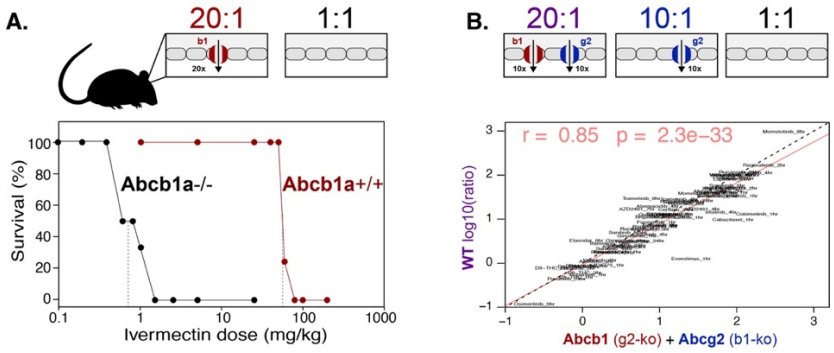

Figure 5. Additive resistance mechanisms create redundancy that blunts the efficacy of MDR inhibitors. (A) Early enthusiasm for MDR (multidrug resistance) inhibitors stemmed from dramatic preclinical findings—such as the first study of ABC-transporter knockouts in mice—where deletion of Abcb1a, the murine ortholog of human ABCB1, caused >20-fold increases in brain exposure to ivermectin and corresponding CNS toxicity.[39] These striking results suggested that inhibition of single drug-efflux pumps could reverse tissue-level drug resistance. **(B)** However, a comprehensive analysis of 76 clinically approved drugs across four murine models (WT, Abcb1a$^{-/-}$, Abcg2$^{-/-}$, and double knockouts) reveals a more nuanced reality. For most drugs, single knockouts produced only modest changes in tissue-to-plasma ratios, indicating that resistance is distributed across multiple enzymes.[40] Our additive resistance model—based on individual barrier ratios—accurately predicted these outcomes, showing that even large resistance effects (e.g., 20×) are functionally diluted when shared between multiple transporters. These findings explain the limited clinical success of MDR inhibitors and highlight the challenge of overcoming redundant, additive resistance networks in cancer and barrier tissues.

These data validate the central prediction of our model: resistance mechanisms act in parallel, and redundancy blunts the effect of targeting any single transporter. Dramatic single-gene phenotypes, such as those seen with ivermectin, are exceptions rather than the rule.[39] In summary, Figure 5 shows that transporter redundancy explains the limited clinical impact of MDR inhibitors while validating the additive resistance predicted by our parallel flux model.[40] Crucially, it also suggests that MDR inhibition may still hold value for the minority of drugs controlled by a single dominant pathway.

DISCUSSION:

This study establishes a practical bridge between laboratory potency metrics and clinical efficacy by normalizing *in vitro* drug sensitivity to clinically achievable exposure. Using the NCI1970-Meta dataset—spanning 49,002 patients, 30 chemotherapies, and 18 cancer types—we show that raw dose-response metrics are poor predictors of patient response. However, when potency is expressed relative to exposure (C_{max} or TR_{min}), translational predictiveness is restored, reconciling laboratory and clinical drug rankings. This provides a simple but powerful framework for integrating pharmacology into preclinical screening.

Key insights for laboratory operations. Our analysis produced two practical recommendations that labs can implement immediately to improve biomarker discovery and drug-repurposing studies:

- Screen the full clinical range (0 – C_{max}).
    - Why it matters: Limiting assays to 10 μM misses antimetabolites and electrophiles that achieve 100–300 μM in patients, leading to costly false negatives and unreliable AI training data.
    - Action: Extend dose ranges to the unbound C_{max} to capture clinically relevant activity and align lab measurements with what oncologists actually deliver.[10]
- Normalize potency by exposure (IC_{50} / C_{max}).
    - Why it matters: In ***single-drug*** biomarker studies, raw AUC/IC_{50} values are sufficient because comparisons happen within the same compound. But in ***multi-drug*** drug-prioritization studies, raw AUC/IC_{50}s are not comparable across agents with different clinical exposure limits.[6]
    - Action: Report and rank candidates by IC_{50} / C_{max}, creating a unitless metric that reflects clinical feasibility and avoids prioritizing drugs with low achievable plasma levels.[43]

This exposure-adjusted approach already underpins the drug-drug interaction (DDI) field, where investigators at the Food and Drug Administration (FDA) have recommend dividing an inhibitor's *in vitro* IC_{50} by its unbound C_{max} to flag compounds likely to cause CYP or transporter inhibition:[43]

$$\frac{IC_{50}}{C_{max}} < 1 \qquad (1)$$

The resulting ratio is unitless, comparable across drugs, and interpreted against fixed thresholds. Our findings extend this framework to oncology, demonstrating that exposure-normalized potency corrects mis-ranked chemotherapies and offers a rigorous tool for drug-repurposing screens[6]—reinforcing the broader utility of this metric in clinical pharmacology.

The implications of this framework are broad. For functional precision medicine platforms—such as organoids, xenografts, and high-content phenotypic assays—exposure normalization ensures that candidate therapies are prioritized based on achievable pharmacology rather than misleading IC_{50} cutoffs. For AI-driven drug prediction models, which are often trained on raw IC_{50} or AUC data, incorporating exposure normalization can prevent systematic biases that undervalue antimetabolites or overvalue agents like gemcitabine. In addition to NCI1970-Meta's ORR data, our work has digitized two recently published resources on TR across 1000 clinical drugs and xenobiotics[18] and $C_{max,u}$ data across 145 cancer drugs.[10]

Biomarker discovery provides an important contrast. Unlike cross-drug comparisons, biomarker analyses typically evaluate multiple patients treated with the same drug-regimen. Because drug exposure is constant within these studies, they are largely robust to the laboratory–clinical dichotomy that confounds potency metrics across drugs. Indeed, we confirm general concordance of 234 cytotoxic chemotherapy biomarkers across both clinical (TCGA) and *in vitro* (CCLE) datasets. Our results show that the collective biomarker literature retains predictive value, particularly for metabolic and transporter pathways, and is less vulnerable to the distortions revealed by potency analyses.

Limitations of the study. The NCI1970-Meta dataset draws primarily from trials conducted between 1955 and 2000, raising questions about its relevance to modern oncology. However, these single-agent studies represent the only large-scale clinical record of cytotoxic monotherapies, as most contemporary trials investigate drug combinations. We curated and reconciled the dataset to maximize transparency, but recognize that heterogeneity in trial design, reporting standards, and patient populations remains a constraint. Likewise, the pharmacokinetic benchmarks (Cmax, TRmin) were curated from therapeutic drug monitoring

studies and reflect population averages;[10,18] interpatient variability and formulation differences introduce additional uncertainty. Finally, while our framework improves translational fidelity at the level of potency metrics, clinical efficacy is shaped by many factors—including tumor heterogeneity, immune interactions, and host toxicity—that extend beyond drug exposure alone. We present exposure normalization not as a complete predictor of clinical outcomes but as a *first-order correction* that restores utility to preclinical pharmacology data.

Looking forward, the NCI1970-Meta dataset provides the first harmonized benchmark for linking single-agent trial outcomes to laboratory drug screens. By unifying legacy clinical data with modern pharmacogenomic resources, it enables systematic evaluation of new metrics, biomarkers, and AI frameworks against real-world patient outcomes. Beyond chemotherapy, the same principles can guide the development of targeted therapies and immunotherapies, where integrating exposure and resistance mechanisms will be equally critical.

In conclusion, exposure-normalized potency represents a practical and generalizable advance for translational oncology. By screening to C_{max} and normalizing dose-response metric, laboratories and computational models alike can align more faithfully with clinical outcomes. This framework clarifies historical discrepancies while reaffirming the robustness of biomarker studies, together charting a path toward more predictive functional assays, biomarker panels, and AI models, and ultimately accelerating the development of effective new therapies.

ACKNOWLEDGEMENTS

E.F.D. acknowledges Matthew Clarke for helpful discussion and the University of Georgia for financial support.

AUTHOR CONTRIBUTIONS

Role	Initials
Conceptualization	EFD
Data curation	EFD, TP, EH
Formal Analysis	EFD,LH,EH, TP, CJ
Investigation	EFD,LH,EH,TP
Methodology	EFD
Project administration	EFD
Supervision	EFD
Visualization	EFD
Writing – original draft	EFD
Writing – review & editing	EFD,LH,EH,TP

DECLARATION OF INTERESTS

The authors declare no conflicts of interest.

METHODS:

Computational Method Details

CURATION OF SINGLE-AGENT RESPONSE RATE DATA (Figure S1-3 and Figure 2A)

Our single-agent dataset (Figure 1A) compiles overall response rates (ORR) of first line treatment of metastatic cancer therapies from phase 2 trials published between 1950 and 2000. The source data was assembled from a textbook and 48 meta-reviews published between 1970 and 2000 (Figure S1). Final data is available in excel format in Supp Data 01 and each stage of curation is documented in Supp Table 02. A description of our curation procedure is provided below and visualized graphically in Figures S01-03.

Briefly, raw response numbers were manually encoded from the NCI textbook: *Single Agents in Cancer Chemotherapy*[9] as a 4 column matrix (Figure S1A) in Excel (Supp Data 02 – NCI1970-update > NCI 1955-70). Once NCI1970 was encoded, we converted it into matrix format to assess the number of gaps (58%) and attempt initial analyses (Figure S1B). Unfortunately, the high sparsity (58%) prevented analysis of sufficient quality and so to reduce this sparsity we: (1) consolidated all lymphoma, leukemia and sarcoma data and added data from an addition 48 meta-reviews published between 1970 and 2000 (Figure S1C) and referenced in the supporting information.

These meta-reviews were compiled in a similar format (Figure S2A) which were focused on individual drugs (Supp Table 02 – NCI1970-update > drug-meta) or individual cancers (Supp Data 02 – NCI1970-update > NCI cancer-meta). 15 of these reviews were published by the same authors as the NCI1970 textbook and followed similar study inclusion criteria. Our main criteria for including meta-review by other authors were: (1) Phase 2 single-agent data (2) WHO definitions of overall response rates and (3) first line treatment of metastatic disease. To prevent inclusion of duplicate data, we designed an algorithm to parse our compiled meta-reviews and update the original NCI1970 data set with new data, whenever more statistically powered datasets were available (Figure S2B). Briefly for each drug-cancer pair, we identified the largest compiled data set and if that data set was larger than the NCI-dataset (in terms of total evaluable patients), we replaced the updated total

number of patients and total responders (Figure S2B). In this manner we were able fill many gaps in NCI1970 and add an additional 25 drugs yielding an updated dataset, NCI1970-meta, which had a significantly reduced number of gaps in the data set (NCI1970-meta: 41% sparse vs NCI1970: 58% sparse; Figure S3). Finally, individual response rates were calculated and cross-referenced with response rates reported in oncology textbooks to confirm concordance (e.g. pages 548, 637, 748, 783, 823, 858, 1035, 1065, 1192, 1305, 1431, 1550 in the 11th edition of *Cancer - Principles & Practice of Oncology*).[26]

CURATION OF FDA-INDICATION DATA (Figure 2C)

To capture "standard practice information" for our 34 drugs and 19 cancers, we manually encoded FDA-indications listed in the 2015 Physicians Cancer Chemotherapy Drug Manual (Figure 2C).[14] First, we constructed a "blank matrix" excel file for overlapping drugs and cancers and entered "0" for each cancer-drug pair. Next, we manually encoded all clinical indications in this manual by replacing each "0" with "1" for cancer-type for which each drug was indicated (Supp Data 01 – Figure Data > FDA (Figure 2B)).

IN VITRO DRUG SENSITIVITY DATASETS (Figure 2E, S4-5)

All *in vitro* drug sensitivity data was obtained from the Dependency Map Data Portal (https://depmap.org/portal/download/) release 20Q2. Cancer Therapeutic Response Portal (CTRPv2)[2] and Genomics of Drug-Sensitivity in Cancer (GDSC)[3] from the Broad and Sanger institutes, respectively, were used as they are the most studied drug-sensitivity datasets currently available. For both CTRPv2 and GDSC, raw Area Under the Curve (AUC) metrics were chosen to quantify overall drug-resistance due to its availability across all datasets. Next, all CTRPv2 and GDSC AUC's were normalized to 1 and converted to drug-sensitivity by subtraction from 1 (1-AUC). Normalization of drug-drug variation in drug-sensitivity (1-AUC) was performed by calculating the z-score's across each row/drug:

$$Z = \frac{x_{drug} - \mu_{drug}}{\sigma_{drug}}$$

Where x_{drug} is 1-AUC for each cell-line, μ_{drug} is the mean across all cell-lines, $\sigma_{drug}$is the standard deviation across all cell-lines. Finally, GDSC and CTRPv2 was visualized as lineage-average across raw (Figure S4A-B) and drug-normalized 1-AUC (Figure S4C-D) to enable comparison with clinical cancer-type response rates (Supp Data 01). All data processing steps were conducted using R statistical software (R Project for Statistical Computing RRID:SCR_001905; https://www.r-project.org).

BENCHMARKING OF INVITRO VS CLINCIAL DATA (Figure 3):

To evaluate the translational relevance of *in vitro* drug-response metrics, we compared PRISM-derived area under the curve (AUC) values to clinical overall response rates (ORRs) from our NCI1970-meta dataset. For each drug–cancer pair, we averaged AUC and ORR values across cancer types and visualized their correspondence using linear regression and annotated scatter plots. To investigate the drug-specific biases in AUC values, we retrieved clinical maximum plasma concentrations (C_{max}) and therapeutic minimum concentrations (TR_{min}) from two recently published pharmacokinetic databases derived from therapeutic drug monitoring studies.[10,18] Unbound C_{max} was calculated using reported protein-binding fractions, and TR_{min} values were converted to molar concentrations using published molecular weights. These exposure metrics were then correlated with AUC values to determine whether clinical dosing constraints could explain discrepancies between *in vitro* and clinical efficacy (Fig 3). All statistical analyses and plots were performed in R and raw code and data is publically available on GitHub (see below).

DRUG-BIOMARKER DEFINITIONS (Figure S6, 4A):

41 drug targets, 47 drug-transporters and 62 drug metabolism genes (Figure S6A-C) were defined for our 34-drug library from the literature curated database DrugBank (RRID:SCR_002700; https://go.drugbank.com/releases/latest).[29,30] An additional 84 biomarkers that indirectly correlate with drug sensitivity were identified from over 200 references catalogued in the supporting information(Figure S6D).

BIOMARKER-EXPRESSION DATASETS (Figure 4A)

All *in vitro* biomarker expression data was obtained from the Cancer Cell-Line Encyclopedia(CCLE) Data Portal (https://depmap.org/portal/download/) release 20Q2.[32,33] All clinical biomarker data was obtained from The Cancer Genome Atlas (TCGA) PanCancer[31] Data-Release at cBioPortal (https://www.cbioportal.org/datasets). For both datasets, RNA sequencing (RNAseq) was used to quantify biomarker expression in transcripts per million (TPM) and transformed to log scale (log2(TPM+1)).

BIOMARKER SET CORRELATION ANALYSIS (Figure 4)

The predictive power of the set of all experimental biomarkers was assessed by summing the mRNA expression of all biomarkers (Figure S6) across lineage averaged RNAseq data sets (Figure 2A) and assessing the correlation with clinical response data (Figure 2B). First, a single biomarker matrix was constructed by combining drug-targets, drug metabolism, drug-transport and indirect target genes and assigning + or – signs to each drug-gene pair based on the biomarkers association with chemo-sensitivity or chemo-resistance, respectively. Direct and indirect drug-targets were assigned a positive sign(Figure S6A,D), drug metabolism genes were assigned a negative sign (Figure 6C) and drug transport genes were assigned a positive sign if they were influx transporters and a negative sign if they were efflux transporters (Figure 6B). Second, biomarker expression data from CCLE and TCGA was converted to TPM to better reflect the log-normal distribution of enzyme expression levels ([E]t). Third, the biomarker identity matrix was multiplied by the biomarker-expression matrix to produce a cancer drug prediction matrix (Figure 2A). Fourth, drug-wise correlations between this prediction matrix (Figure 2A) and the clinical response matrix (Figure 2B) were calculated and plotted in Figure 1C-D. Color coding of each correlation barplot was in proportion to the relative correlation coefficient of each biomarker submatrix in Figure S6 to visualize the contribution of metabolism, transport and target and indirect target genes.

INDIVIUDAL BIOMARKER CORRELATION ANALYSIS (Figure S7-S9).

As the taxanes, anthracyclines and vinca alkaloids showed some of the highest correlation (Figure 2C-D) we next assessed the correlation of individual biomarkers across these classes of compounds as illustrated in Figure S7. Clinical and *in vitro* correlation coefficients were plotted on the x and y-axes, respectively to enable direct comparison of the most predictive biomarkers for vinca alkaloids, taxanes and anthracyclines in Figure S8. Finally, multiple linear regressions were performed on across all biomarkers for these three drug-classes to demonstrate the independent correlation of these biomarkers with drug-sensitivity variance (Figure S9).

LITERATURE EXPERIMENTAL DATA ANALYSIS (Figure S10)

To assess the agreement of our chemo-resistance model (equation 10) with historical experimental data on the anthracyclines, we manually extracted raw drug-uptake and dose-response data from Roovers, et al. Leuk Res 1999, 23, 539-548 into excel (Supp Data 03). Drug uptake data was fit to simple exponential transfer morel ($[drug]_{intra} = [drug]_{ss} * (1 – e^{-k*time})$) while dose response curves were fit to a simplified Langmuir-Hill model (fraction viability = $IC_{50} / (IC_{50} + [drug])$)

GENERAL DATA ANALYSAIS & VISUALIZATION:

All data processing, analysis and visualization in this manuscript was conducted using R statistical software (https://www.r-project.org). Linear regressions and multiple linear regressions were conducted using the lm and glm functions in R; while Heatmaps were created using the heatmap.2 function in the gplots package (https://cran.r-project.org/web/packages/gplots/index.html) and the colorRampPalette function from the RColorBrewer package (RRID:SCR_016697; https://cran.r-project.org/web/packages/RColorBrewer/index.html).

Resource Availability

***Lead Contact**:* Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Eugene Douglass (eugene.douglass@uga.edu).

***Materials Availability**:* This study did not generate new unique reagents.

***Data and Code Availability**:* Data visualization and analysis code is publicly available through a GitHub Repository (https://github.com/fingolfn/nci1970).

REFERENCES:

1 Shoemaker, R. H. The NCI60 human tumour cell line anticancer drug screen. *Nat Rev Cancer* **6**, 813–823 (2006). https://doi.org/10.1038/nrc1951

2 Basu, A. *et al.* An Interactive Resource to Identify Cancer Genetic and Lineage Dependencies Targeted by Small Molecules. *Cell* **154**, 1151–1161 (2013). https://doi.org/10.1016/j.cell.2013.08.003

3 Iorio, F. *et al.* A Landscape of Pharmacogenomic Interactions in Cancer. *Cell* **166**, 740–754 (2016). https://doi.org/10.1016/j.cell.2016.06.017

4 Corsello, S. M. *et al.* Discovering the anti-cancer potential of non-oncology drugs by systematic viability profiling. *Nat Cancer* **1**, 235–248 (2020). https://doi.org/10.1038/s43018-019-0018-6

5 Aykan, N. F. & Ozatli, T. Objective response rate assessment in oncology: Current situation and future expectations. *World J Clin Oncol* **11**, 53–73 (2020). https://doi.org/10.5306/wjco.v11.i2.53

6 Chen, Y. *et al.* Optimizing drug sensitivity assays in patient-derived tumor organoids: a comparison of IC50 estimation methods and experimental parameters. *Biol Methods Protoc* **10**, bpaf012 (2025). https://doi.org/10.1093/biomethods/bpaf012

7 Griffin, R. J., Avery, E. & Xia, C. Q. Predicting Approximate Clinically Effective Doses in Oncology Using Preclinical Efficacy and Body Surface Area Conversion: A Retrospective Analysis. *Front Pharmacol* **13**, 830972 (2022). https://doi.org/10.3389/fphar.2022.830972

8 Graham, M. A. & Workman, P. The impact of pharmacokinetically guided dose escalation strategies in phase I clinical trials: critical evaluation and recommendations for future studies. *Ann Oncol* **3**, 339–347 (1992). https://doi.org/10.1093/oxfordjournals.annonc.a058203

9 Livingston, R. B. & Carter, S. K. *Single agents in cancer chemotherapy*. (IFI/Plenum, 1970).

10 Liston, D. R. & Davis, M. Clinically Relevant Concentrations of Anticancer Drugs: A Guide for Nonclinical Studies. *Clin Cancer Res* **23**, 3489–3498 (2017). https://doi.org/10.1158/1078-0432.CCR-16-3083

11 Adam, G. *et al.* Machine learning approaches to drug response prediction: challenges and recent progress. *NPJ Precis Oncol* **4**, 19 (2020). https://doi.org/10.1038/s41698-020-0122-1

12 Ali, M. & Aittokallio, T. Machine learning and feature selection for drug response prediction in precision oncology applications. *Biophys Rev* **11**, 31–39 (2019). https://doi.org/10.1007/s12551-018-0446-z

13 Wells, J. D., Griffin, J. R. & Miller, T. W. Pan-Cancer Transcriptional Models Predicting Chemosensitivity in Human Tumors. *Cancer Inform* **20**, 11769351211002494 (2021). https://doi.org/10.1177/11769351211002494

14 Chu, E. & DeVita, V. T. *Physicians' Cancer Chemotherapy Drug Manual 2015*. (Jones & Bartlett Learning, 2015).

15 Choi, J. H. *et al.* Comparison of WHO and RECIST criteria for response in metastatic colorectal carcinoma. *Cancer Res Treat* **37**, 290–293 (2005). https://doi.org/10.4143/crt.2005.37.5.290

16 Park, J. O. *et al.* Measuring response in solid tumors: comparison of RECIST and WHO response criteria. *Jpn J Clin Oncol* **33**, 533–537 (2003). https://doi.org/10.1093/jjco/hyg093

17 Howlader, N. *et al. SEER Cancer Statistics Review, 1975-2017*. (National Cancer Institute, 2019).

18 Schulz, M., Iwersen-Bergmann, S., Andresen, H. & Schmoldt, A. Therapeutic and toxic blood concentrations of nearly 1,000 drugs and other xenobiotics. *Crit Care* **16**, R136 (2012). https://doi.org/10.1186/cc11441

19 Twomey, J. D., Brahme, N. N. & Zhang, B. Drug-biomarker co-development in oncology - 20 years and counting. *Drug Resist Updat* **30**, 48–62 (2017). https://doi.org/10.1016/j.drup.2017.02.002

20 Schwaederle, M. *et al.* Association of Biomarker-Based Treatment Strategies With Response Rates and Progression-Free Survival in Refractory Malignant Neoplasms: A Meta-analysis. *JAMA Oncol* **2**, 1452–1459 (2016). https://doi.org/10.1001/jamaoncol.2016.2129

21 Goossens, N., Nakagawa, S., Sun, X. & Hoshida, Y. Cancer biomarker discovery and validation. *Transl Cancer Res* **4**, 256–269 (2015). https://doi.org/10.3978/j.issn.2218-676X.2015.06.04

22 Merrill, N. M. *et al.* Integrative Drug Screening and Multiomic Characterization of Patient-derived Bladder Cancer Organoids Reveal Novel Molecular Correlates of Gemcitabine Response. *Eur Urol* **86**, 434–444 (2024). https://doi.org/10.1016/j.eururo.2024.05.026
23 Dahl, M. J. *et al.* Gemcitabine combination therapies induce apoptosis in uterine carcinosarcoma patient-derived organoids. *Front Oncol* **14**, 1368592 (2024). https://doi.org/10.3389/fonc.2024.1368592
24 Boileve, A. *et al.* Organoids for Functional Precision Medicine in Advanced Pancreatic Cancer. *Gastroenterology* **167**, 961–976 e913 (2024). https://doi.org/10.1053/j.gastro.2024.05.032
25 Cartry, J. *et al.* Implementing patient derived organoids in functional precision medicine for patients with advanced colorectal cancer. *J Exp Clin Cancer Res* **42**, 281 (2023). https://doi.org/10.1186/s13046-023-02853-4
26 DeVita, V. T., Lawrence, T. S. & Rosenberg, S. A. *Cancer - Principles & Practice of Oncology, 11th Ed*. (Wolters Kluwer, 2015).
27 Editors. A call for pragmatism in cancer research. *Nat Rev Clin Oncol* **15**, 193 (2018). https://doi.org/10.1038/nrclinonc.2018.41
28 Olaussen, K. A. & Postel-Vinay, S. Predictors of chemotherapy efficacy in non-small-cell lung cancer: a challenging landscape. *Ann Oncol* **27**, 2004–2016 (2016). https://doi.org/10.1093/annonc/mdw321
29 Wishart, D. S. *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* **46**, D1074–D1082 (2018). https://doi.org/10.1093/nar/gkx1037
30 Wishart, D. S. *et al.* DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res* **34**, D668–672 (2006). https://doi.org/10.1093/nar/gkj067
31 Weinstein, J. N. *et al.* The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* **45**, 1113–1120 (2013). https://doi.org/10.1038/ng.2764
32 Ghandi, M. *et al.* Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–+ (2019). https://doi.org/10.1038/s41586-019-1186-3
33 Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012). https://doi.org/10.1038/nature11003
34 Leonessa, F. & Clarke, R. ATP binding cassette transporters and drug resistance in breast cancer. *Endocr Relat Cancer* **10**, 43–73 (2003). https://doi.org/10.1677/erc.0.0100043
35 Segel, I. H. *Enzyme kinetics : behavior and analysis of rapid equilibrium and steady state enzyme systems*. (Wiley, 1975).
36 Bamford, C. H., Tipper, C. F. H. & Compton, R. G. *Comprehensive chemical kinetics*. (Elsevier Pub. Co., 1969).
37 Michaelis, V. L. & Menten, M. M. L. Die kinetik der invertinwirkung (The Kinetics of Invertase Action). *Biochemische Zeitschrift* **49**, 333–369 (1913).
38 Missner, A. & Pohl, P. 110 years of the Meyer-Overton rule: predicting membrane permeability of gases and other small compounds. *Chemphyschem* **10**, 1405–1414 (2009). https://doi.org/10.1002/cphc.200900270
39 Schinkel, A. H. *et al.* Disruption of the mouse mdr1a P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs. *Cell* **77**, 491–502 (1994). https://doi.org/10.1016/0092-8674(94)90212-7
40 Hughes, E. A., Zielinski, R., Ray, A. E., Priebe, W. & Douglass Junior, E. F. Development of a Novel Tool To Demystify Drug Distribution at Tissue-Blood Barriers. *Chembiochem*, e202200804 (2023). https://doi.org/10.1002/cbic.202200804
41 Roovers, D. J., van Vliet, M., Bloem, A. C. & Lokhorst, H. M. Idarubicin overcomes P-glycoprotein-related multidrug resistance: comparison with doxorubicin and daunorubicin in human multiple myeloma cell lines. *Leuk Res* **23**, 539–548 (1999). https://doi.org/10.1016/s0145-2126(99)00041-7

42 Libby, E. & Hromas, R. Dismounting the MDR horse. *Blood* **116**, 4037–4038 (2010). https://doi.org/10.1182/blood-2010-09-304311

43 Zhang, L., Zhang, Y. D., Zhao, P. & Huang, S. M. Predicting drug-drug interactions: an FDA perspective. *AAPS J* **11**, 300–306 (2009). https://doi.org/10.1208/s12248-009-9106-3