https://doi.org/10.31925/farmacia.2026.4.21 **ORIGINAL ARTICLE**

COMPARATIVE ANALYSIS OF DRUG RESPONSE METRICS ACROSS TUMOUR TYPES UNDER CAMPTOTHECIN EXPOSURE

LINA CHUB [1]*, WASHINGTON BARBOSA DA SILVA [2], FELIPE DE OLIVEIRA LOPES CAVALCANTI [3], RONNI GERALDO GOMES DE AMORIM [1]

[1]Faculty of Science and Technology in Engineering, University of Brasília (UnB), Brasília, Brazil
[2]Physics Teaching Division, Military College of Brasília, Brasília, Brazil
[3]Department of Social Medicine, Faculty of Medicine, University of Brasília (UnB), Brasília, Brazil

**corresponding author: chublina0@gmail.com*

Manuscript received: April 2026

Abstract

The LN(IC_{50}) and AUC metrics are standard in cancer pharmacogenomics, but their interchangeability remains poorly understood. We tested the hypothesis that a high correlation between these metrics does not necessarily imply identical results in the binary classification of tumour cell line sensitivity. Using GDSC2 data from 116 cell lines (BLCA, BRCA, CESC, GBM), a correlation analysis was performed, and the frequency of discordance between LN(IC_{50}) and AUC for camptothecin was evaluated. Robustness was assessed by varying classification thresholds and comparing RMSE between concordant and discordant observations. Despite a high correlation ($\rho = 0.967$), discordance reached 8.6%. Alternative thresholds reduced discordance to 6.9%. RMSE differences were not significant ($p = 0.0808$), although the discordant group was small. Discordant observations clustered near the classification boundaries were observed ($p < 0.001$), which is to be expected when two highly correlated measures are divided at the median. The results demonstrate that highly correlated continuous metrics do not guarantee that binary outcomes will be identical, and emphasise the value of jointly analysing several metrics of drug response when interpreting borderline cases.

Rezumat

Indicatorii LN(IC_{50}) și AUC sunt standard în farmacogenomica oncologică, însă interschimbabilitatea lor rămâne insuficient înțeleasă. Am testat ipoteza conform căreia o corelație ridicată între acești indicatori nu conduce, în mod necesar, la rezultate identice în clasificarea binară a sensibilității liniilor celulare tumorale. Utilizând datele GDSC2 pentru 116 linii celulare (BLCA, BRCA, CESC, GBM), s-a efectuat o analiză de corelație și s-a evaluat frecvența discordanței dintre LN(IC_{50}) și AUC pentru camptotecină. Robustețea a fost evaluată prin variația pragurilor de clasificare și prin compararea RMSE între observațiile concordante și cele discordante. Deși a fost observată o corelație ridicată ($\rho = 0,967$), discordanța a atins 8,6%. Pragurile alternative au redus discordanța la 6,9%. Diferențele de RMSE nu au fost semnificative ($p = 0,0808$), deși grupul discordant a fost restrâns. De asemenea, observațiile discordante s-au grupat în apropierea limitelor de clasificare ($p < 0,001$), fapt previzibil atunci când două măsuri puternic corelate sunt divizate la mediană. Rezultatele demonstrează că indicatorii continui puternic corelați nu garantează obținerea unor rezultate binare identice și subliniază utilitatea analizării conjuncte a mai multor indicatori de răspuns la tratament în interpretarea cazurilor limită.

Keywords: area under the curve, tumour cell line, dose-response relationship, half maximal inhibitory concentration, pharmacogenomics

Introduction

The development of modern personalised medicine and precision oncology in particular relies on the use of high-throughput pharmacogenomic platforms (1, 2). Resources such as the Genomics of Drug Sensitivity in Cancer (GDSC) project provide a standardised basis for the comparative analysis of drug efficacy (3). GDSC data allow assessment of tumour model sensitivity to tested compounds. More reliable analytical tools are required for in-depth and accurate interpretation of quantitative indicators of drug response. The widespread use of these measures does not negate concerns about their sensitivity to the dataset's structure. Insufficient attention has been paid to the impact of the choice of cut-off point on the stability of the classification. The question of the distribution of observations near the cut-off threshold remains open, despite the fact that this factor is critical when assessing the reliability of the classification (4). Despite the strong correlation between key metrics widely used in pharmacology, toxicology and biomedicine (LN(IC_{50}) and AUC), the consistency of binary classification results obtained using these metrics remains a matter of debate. LN(IC_{50}) and AUC characterise different aspects of the dose-response curve, which could potentially lead to different classifications of the same observations (5, 6). This factor can lead to differing interpretations of the sensitivity of the same tested cell lines. We investigate whether a high correlation between these metrics is sufficient for the reliable identification of

sensitive and resistant phenotypes, or whether it may mask underlying discrepancies in their classification. To rule out inter-drug variability, the study design focused on analysing the effect of a single drug (camptothecin) on four tumour types (BLCA, BRCA, CESC, GBM). The focus of this study was on a comparative analysis of the binary classification of tumour lines using the LN(IC_{50}) and AUC metrics, an analysis of quantitative differences between groups of concordant and discordant observations, and an assessment of the robustness of the identified discordant observations when the threshold values for binary classification were altered. The results obtained may contribute to improving the interpretation of pharmacogenomic study results and the assessment of the robustness of sensitivity classification in tumour models.

Materials and Methods

Dataset

The study is based on the open-access dataset from the Genomics of Drug Sensitivity in Cancer (GDSC2) project (2, 3). A precompiled snapshot of the GDSC2 dataset was downloaded from Kaggle on 13 August 2024. This dataset contains precomputed drug response metrics derived from the GDSC2 pharmacogenomic resource and was used as the data source for all analyses performed in this study. The analyses presented in this study are based exclusively on this dataset snapshot to ensure reproducibility of the reported results. Camptothecin (GDSC2: DRUG_ID = 1003, DRUG_NAME = Camptothecin, PUTATIVE_TARGET = TOP1, PATHWAY_NAME = DNA replication), a topoisomerase I inhibitor, was selected as the model agent. Drug response data were extracted from the GDSC2 dataset snapshot described above. Its classical and well-studied mechanism of action, as well as the wide range of responses observed in the tested models, provide sufficient variability in the data for the purposes of this study. Only records containing both LN(IC_{50}) and AUC values were included, allowing pairwise comparison of both response metrics for each cell line. Cell lines from the following tumours were selected: urothelial carcinoma of the bladder (BLCA), invasive breast carcinoma (BRCA), squamous cell carcinoma of the cervix (CESC), and glioblastoma multiforme (GBM). Tumour groups were assigned according to the TCGA_DESC annotation provided in the dataset. After filtering, the final sample size was 116 cell lines. A single global dataset-wide filter was applied, retaining only observations with complete values for LN(IC_{50}) and AUC corresponding to the selected tumour groups. No additional processing of duplicate records, replicate screens or curve-quality filtering was performed. The data used are publicly available and de-identified; no additional ethical approval was required. RMSE values were obtained directly from the precompiled GDSC2 dataset snapshot used in this study. No additional curve fitting or RMSE calculation was performed in the present study.

Response metrics and classification

The pharmacological response was assessed using the natural logarithm of the half-maximal inhibitory concentration (LN(IC_{50})) and the area under the dose–response curve (AUC) (5, 6). The study utilised the AUC metric provided in the GDSC2 dataset; lower AUC values corresponded to greater sensitivity of the cell lines to the drug. Both metrics have the same interpretative direction: low values correspond to high sensitivity, whilst higher values indicate a more resistant response. To assess the performance of the binary classification based on LN(IC_{50}) and AUC, a binary classification was applied using a single median threshold calculated across the entire combined sample (n = 116). The choice of the median was dictated by the need for a statistically neutral division of continuous data into comparable groups, whilst minimising the influence of distribution asymmetry and artefacts. This approach was implemented to determine the degree of discordance in the binary classification of the sensitivity of the same cell lines. A classification was considered concordant if both metrics assigned the cell line to the same category ('sensitive' or 'resistant'). Discordance was defined as the opposite binary classification of the same cell line when using LN(IC_{50}) and AUC. The overall proportion of discordant observations was calculated both for the entire sample and separately for each tumour group under investigation.

Statistical analysis

The statistical analysis aimed to describe the distributions of the LN(IC_{50}) and AUC metrics across four tumour groups (BLCA, BRCA, CESC, GBM) after exposure to camptothecin. An assessment of intergroup differences was also carried out, along with an analysis of the association between the metrics and their consistency. The calculations were carried out in Python (3.12.7) using the pandas (2.2.2), NumPy (1.26.4), SciPy (1.13.1) and Matplotlib (3.9.2) libraries. The analysis was conducted exclusively on curated GDSC2 data. Reproducibility was ensured through fixed algorithm parameters. The significance level was set at $\alpha = 0.05$. The choice of robust non-parametric methods was dictated by the specific nature of the asymmetric distribution. All tests were two-tailed. Central tendency was described by the median, and variability by the interquartile range (IQR) (7). Intergroup differences were assessed using the Kruskal–Wallis test (8), followed by pairwise comparisons using the Mann–Whitney test (9) with Holm–Bonferroni correction (10). The Holm–Bonferroni correction was applied only when performing multiple pairwise comparisons between tumour groups. For the remaining independent statistical

analyses (correlation analysis, RMSE analysis, analysis of distances to the median, and sensitivity analysis to threshold selection), no correction for multiple comparisons was applied. The magnitude of the differences was further assessed using the non-parametric effect size measure Cliff's delta (δ) (11), calculated using our own implementation of the algorithm in the Python programming language. Spearman's rank correlation was used to assess the degree of agreement between LN(IC_{50}) and AUC (12). Bootstrap analysis (2,000 repetitions) was applied to verify the robustness of the obtained medians and to assess the stability of the results upon repeated sampling of the data (13). The statistical reliability of the median values was ensured by calculating 95% confidence intervals using the percentile method. The Mann–Whitney non-parametric test was used to compare the distribution of RMSE across the groups (concordant and discordant observations). The RMSE values were obtained directly from the GDSC2 dataset and were used as an indicator of the quality of the dose-response curve approximation for each cell line. No additional RMSE calculations were performed in this study. Analysis using the 40th and 60th percentiles of the distribution was carried out additionally to assess the robustness of the results to the choice of threshold value. The structure of the data distribution was analysed using absolute distances from the median, calculated separately for LN(IC_{50}) and AUC. All statistical analyses were performed on continuous values of LN(IC_{50}) and AUC, with the exception of the concordance analysis, for which a binary classification based on a single median threshold was defined in advance.

Results and Discussion

Sample structure and distribution of tumour types

The study covered a sample of 116 cell lines. The structure of this sample was characterised by asymmetry, and the distribution of models was uneven, which is a characteristic feature of database creation. The uneven representation of the cell lines studied is due to the varying frequency of tumour types in the initial dataset. Overall, BRCA cell lines predominated (43.1%), whilst cervical squamous cell carcinoma (CESC) lines accounted for 12.1% (Table I).

Table I

Sample distribution by tumour type

Tumour type	n	% of total
BLCA	18	15.5
BRCA	50	43.1
CESC	14	12.1
GBM	34	29.3
Total	**116**	**100.0**

Tumour-specific sensitivity patterns to camptothecin

The analysis includes only observations with available LN(IC_{50}) and AUC values. The BLCA and CESC cell lines were the most sensitive to camptothecin. The median LN(IC_{50}) values for these tumour groups were below -2.6, indicating their high susceptibility (Table II).

Invasive breast carcinoma (BRCA) cell lines exhibited systemic resistance, as evidenced by the values of both metrics falling within the upper ranges of the distributions. The median LN(IC_{50}) value for this group was -0.984. Statistically significant differences among tumour groups were observed for both LN(IC_{50}) and AUC, as determined by the Kruskal–Wallis test (Table III).

Table II

Distribution of LN(IC_{50}) and AUC by tumour type

Tumour type	n	LN(IC_{50}) median (Q1–Q3)	AUC median (Q1–Q3)
BLCA	18	-2.867 (-3.653 to -1.675)	0.823 (0.751–0.887)
BRCA	50	-0.984 (-2.057 to 0.294)	0.936 (0.885–0.968)
CESC	14	-2.669 (-3.212 to -1.490)	0.833 (0.788–0.891)
GBM	34	-2.392 (-3.361 to -1.493)	0.841 (0.775–0.892)

Values are presented as median (Q1–Q3). Lower values indicate higher sensitivity to camptothecin.

Table III

Intergroup differences in LN(IC_{50}) and AUC

Metric	H statistic (Kruskal–Wallis)	df	p-value
LN(IC_{50})	23.87	3	2.66×10^{-5}
AUC	26.96	3	6.01×10^{-6}

Kruskal–Wallis test results for comparison across four tumour groups (df = 3). Two-sided p-values are shown.

The correlation between LN(IC_{50}) and AUC was assessed under conditions of fixed pharmacological exposure across the four tumour groups tested. The focus was on determining the extent to which metrics of the cytotoxic response in cell lines influence the interpretation of differences between tumour types. The analysis identified reproducible differences in sensitivity to camptothecin (a topoisomerase I inhibitor) among 116 cell lines from the tumour groups under investigation (BLCA, BRCA, CESC, GBM) (1, 14). The robustness of these differences is supported by the application of a range of analytical approaches, including effect-size estimation, distributional analysis, and non-parametric statistical

methods (15). A bootstrap analysis performed at the end of the study verified the results and minimised potential bias in the conclusions arising from random fluctuations within small subsets (*e.g.*, within the CESC group) (13). A key finding was the marked resistance exhibited by cell lines of invasive breast carcinoma (the BRCA tumour group). The observed resistance of BRCA cell lines may reflect biological characteristics previously reported for this tumour type (16, 17). This pattern remained consistent across the statistical approaches applied in the present study. This may be of particular scientific interest. For example, in comparative pharmacogenomics, it could serve as a reliable reference in the context of the search for new molecular indicators of drug response (3).

Distributional shifts and intergroup differences

In Figures 1–2, the distributions for the BRCA group are shifted towards higher values for both the half-maximal inhibitory concentration (LN(IC_{50})) and the area under the curve (AUC).

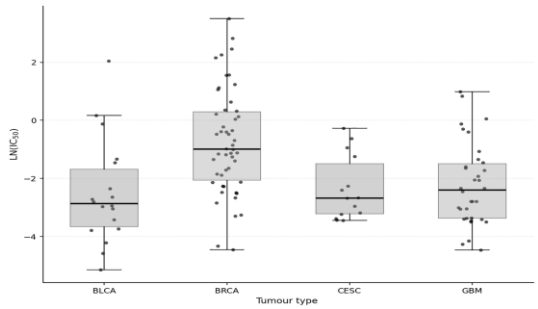

Figure 1.
Distribution of LN(IC_{50}) across tumour groups

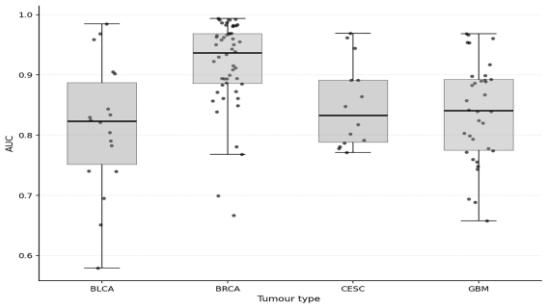

Figure 2.
Distribution of AUC across tumour groups

LN(IC_{50}) values for camptothecin in four tumour groups (BLCA, BRCA, CESC, GBM). Boxes represent the interquartile range (IQR), central lines represent the medians, and whiskers extend to 1.5×IQR. Points represent individual cell lines. Higher LN(IC_{50}) values indicate lower sensitivity.

Area under the dose–response curve (AUC) for camptothecin in four tumour groups (BLCA, BRCA, CESC, GBM). Boxes represent the interquartile range (IQR), central lines represent medians, and whiskers extend to 1.5×IQR. Points represent individual cell lines. Higher AUC values indicate lower sensitivity. Pairwise comparisons of the analysed groups (BLCA, BRCA, CESC, GBM) confirmed the uniqueness of the BRCA group's profile. In all cases, pairs involving BRCA showed higher median differences and large effect sizes (Cliff's $\delta > 0.5$ for both metrics). In comparisons of groups not involving BRCA, the results did not reach statistical significance, indicating that BRCA was the main contributor to the observed intergroup differences (Table IV and Table V). This observation is consistent with previous pharmacogenomic studies reporting variability in drug response across tumour-derived cell lines (18, 19).

Table IV

Pairwise differences in LN(IC_{50}) between tumour groups

Comparison (A *vs* B)	n_1	n_2	$Median_1$	$Median_2$	Median diff	Cliff's δ	p-value (Holm- adjusted)
BRCA vs GBM	50	34	−0.984	−2.392	1.408	0.493	0.0008
BLCA vs BRCA	18	50	−2.867	−0.984	−1.883	−0.564	0.0021
BRCA vs CESC	50	14	−0.984	−2.669	1.685	0.557	0.0063
BLCA vs CESC	18	14	−2.867	−2.669	−0.198	−0.135	1.000
BLCA vs GBM	18	34	−2.867	−2.392	−0.476	−0.144	1.000
CESC vs GBM	14	34	−2.669	−2.392	−0.277	−0.017	1.000

Pairwise comparisons using the Mann–Whitney U test with Holm correction. Effect size is reported as Cliff's δ. Positive δ values indicate higher LN(IC_{50}) in group A

Table V

Pairwise differences in AUC between tumour groups

Comparison (A vs B)	n_1	n_2	$Median_1$	$Median_2$	Median difference	Cliff's δ	p-value (Holm- adjusted)
BRCA vs GBM	50	34	0.936	0.841	0.096	0.564	0.0001
BLCA vs BRCA	18	50	0.823	0.936	−0.113	−0.584	0.0013
BRCA vs CESC	50	14	0.936	0.833	0.104	0.526	0.0115
BLCA vs CESC	18	14	0.823	0.833	−0.009	−0.135	1.000
BLCA vs GBM	18	34	0.823	0.841	−0.018	−0.108	1.000
CESC vs GBM	14	34	0.833	0.841	−0.008	0.059	1.000

Pairwise comparisons using the Mann–Whitney U test with Holm correction. Effect size is reported as Cliff's δ. Positive δ values indicate higher AUC in group A.

Concordance between LN(IC₅₀) and AUC metrics

Spearman's rank correlation coefficient (ρ = 0.967) revealed a strong monotonic relationship between the LN(IC₅₀) and AUC metrics across the combined sample of 116 cell lines, indicating a high degree of rank consistency between the two metrics (Figure 3).

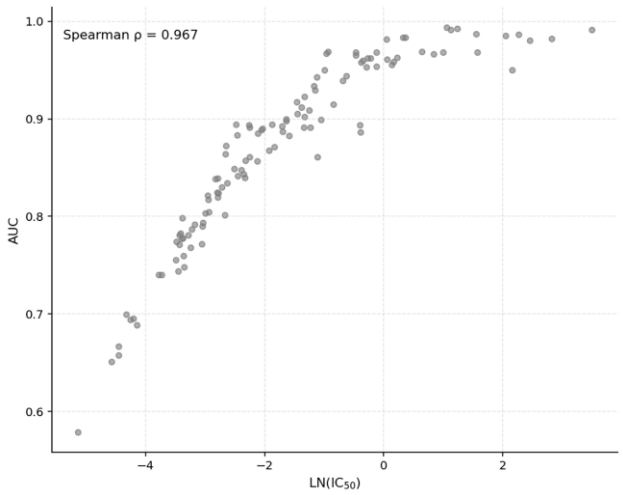

Figure 3.

Monotonic relationship between LN(IC₅₀) and AUC. Scatter plot showing the relationship between LN(IC₅₀) and AUC for all cell lines (n = 116). Spearman's rank correlation coefficient (ρ = 0.967) indicates a strong monotonic association between the metrics

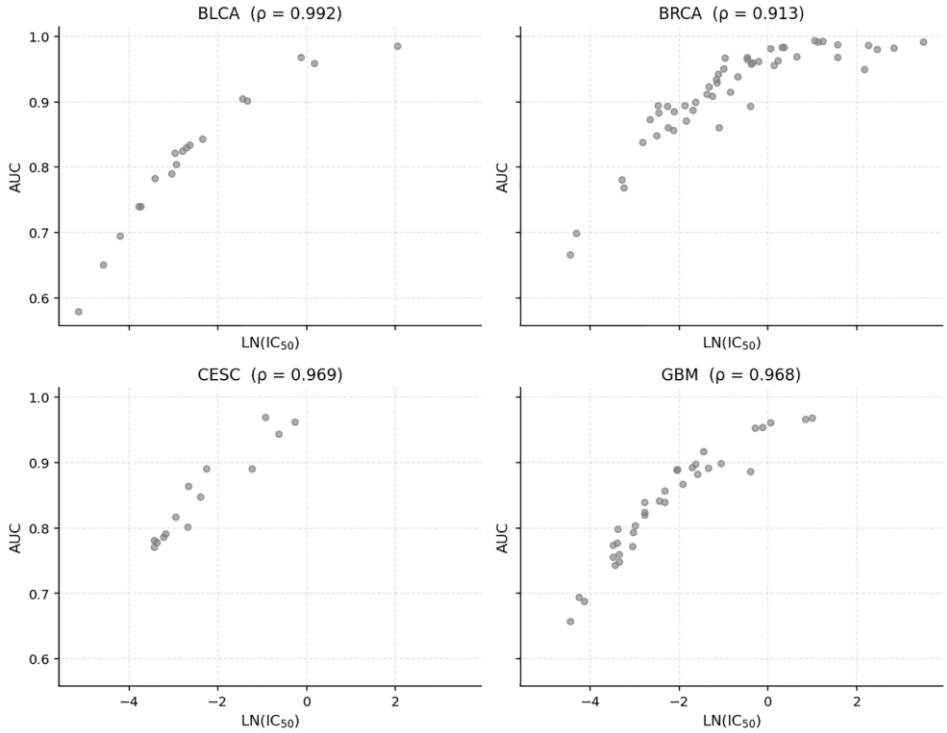

Figure 4.

Monotonic relationship between LN(IC₅₀) and AUC within tumour groups. Scatter plots showing the relationship between LN(IC₅₀) and AUC within each tumour group (BLCA, BRCA, CESC, GBM). Spearman's rank correlation coefficient (ρ) is shown for each group

The same relationship is observed in the scatter plots within each tumour group. Spearman's coefficient ranged from a minimum of ρ = 0.913 for BRCA to a maximum of ρ = 0.992 for BLCA (Figure 4).

An analysis of concordance between the metrics under investigation, despite their high correlation, revealed discordance in 8.6% of cases, corresponding to 10 out of 116 observations (Figure 5).

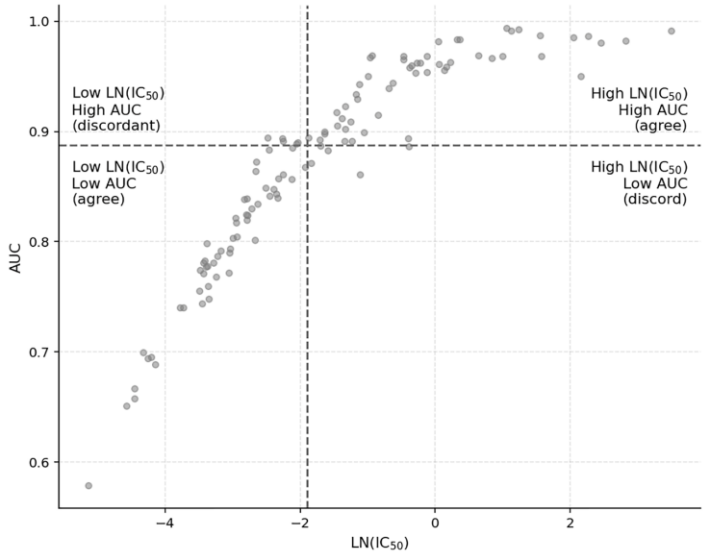

Figure 5.
Consistency of the LN(IC_{50}) and AUC classifications.
Quadrant stratification based on median values of LN(IC_{50}) and AUC. Dotted lines indicate global median thresholds. Observations are classified as consistent when both metrics assign the same sensitivity category and inconsistent otherwise. The overall proportion of inconsistent observations is 8.6%. Regions of discordance indicate observations for which median-based classification differs between LN(IC_{50}) and AUC.

An analysis of the classification based on median thresholds revealed inconsistencies in 8.6% of observations, despite a high correlation ($\rho = 0.97$) as indicated by Spearman's rank correlation between the analysed metrics LN(IC_{50}) and AUC (5, 6). This indicates that even a high correlation does not necessarily guarantee identical binary classification outcomes when LN(IC_{50}) and AUC are used separately (6). Consequently, when using different metrics of drug response, individual observations—particularly those situated close to the median cut-off values—may be classified into different sensitivity categories. These results highlight the need for caution when interpreting binary classifications and confirm the value of considering continuous LN(IC_{50}) and AUC values together when analysing borderline cases (5, 6, 20, 21).

Tumour-specific patterns of discordance

The distribution of inconsistent classifications was specific in nature and depended on tumour type: ranging from a complete absence of discordance in the BLCA group to maximum values of 10.0% and 11.8% in the BRCA and GBM groups, respectively (Table VI), consistent with previous pharmacogenomic studies showing that drug response varies across tumour lineages and biological contexts (20, 22).

Table VI
Frequency of discordant classification based on LN(IC_{50}) and AUC

Tumour type	Discordance (%)
BLCA	0.0
BRCA	10.0
CESC	7.1
GBM	11.8
Overall	8.6

Proportion of observations classified differently by LN(IC_{50}) and AUC using median cut-offs. Discordance is defined as assignment to opposite sensitivity categories.

Extreme sensitivity and resistance phenotypes

Analysis of extreme sensitivity phenotypes (Table VII) revealed that 38% of the BRCA group lines fall in the upper quartile of resistance and only 8% in the lower quartile, further confirming the pronounced resistance of BRCA. No signs of systemic resistance or high sensitivity were identified for GBM cell line metrics. For the BLCA and CESC groups, the distributions shifted towards sensitivity.

Table VII
Extreme sensitivity phenotypes based on LN(IC_{50})

Tumour type	Total (n)	Top-sensitive (n)	Top-sensitive (%)	Bottom-resistant (n)	Bottom-resistant (%)
BLCA	18	6	33.3%	3	16.7%
BRCA	50	4	8%	19	38%
CESC	14	4	28.6%	1	7.1%
GBM	34	5	14.7%	6	17.6%

Distribution of the most sensitive (≤25th percentile) and most resistant (≥75th percentile) cell lines based on LN(IC_{50}).

Robustness of results (bootstrap validation)

To validate the results obtained, a bootstrap analysis (2,000 iterations) was performed. The stability of the calculated median LN(IC_{50}) values in this analysis confirms the identified patterns of sensitivity and verifies the robustness of the results to variations in the sample. The 95% confidence intervals calculated using the percentile method confirm the identified differences in sensitivity among the study groups. The lowest variability in the median was observed among BRCA invasive breast cancer cell lines. The width of the confidence interval for this group was minimal, at 0.985, whereas the CESC group showed the maximum value, CI = 1.989 (Table VIII). The statistical reliability of the results obtained was confirmed using the non-parametric bootstrap method. Bootstrap resampling is widely used to assess the stability of statistical estimates and to construct confidence intervals without relying on strict distributional assumptions (23).

Table VIII

Stability of median LN(IC_{50}) estimates based on bootstrap analysis

Tumour type	n	Median LN(IC_{50})	95% CI	CI width
BLCA	18	-2.867	-3.575 to -1.900	1.675
BRCA	50	-0.984	-1.360 to -0.375	0.985
CESC	14	-2.669	-3.224 to -1.235	1.989
GBM	34	-2.392	-3.041 to -1.712	1.329

Median LN(IC_{50}) values and 95% confidence intervals estimated using the bootstrap method (2,000 bootstrap samples, percentile method). CI width reflects estimate stability.

Influence of curve-fitting quality

A comparison of RMSE values between concordant and discordant cell lines revealed no statistically significant differences (Mann–Whitney U test, p = 0.0808). However, the median RMSE value was slightly higher in the discordant group. The small number of discordant cell lines (n = 10) limits the statistical power of this comparison. Therefore, the results obtained do not allow a reliable conclusion to be drawn regarding the presence or absence of a link between the quality of the dose–response curve approximation and the probability of discordance.

Table IX

Comparison of RMSE values between concordant and discordant cell lines

Group	n	RMSE, median (Q1–Q3)	p-value
Concordant	106	0.0870 (0.0757–0.1100)	
Discordant	10	0.1055 (0.0879–0.1332)	0.0808

Values are presented as median and interquartile range (Q1–Q3). P-value was obtained using the Mann–Whitney U test.

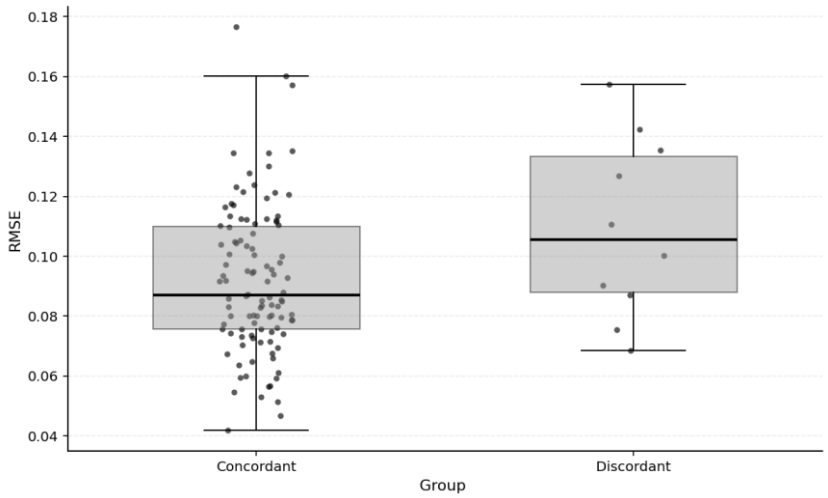

Figure 6.

Distribution of RMSE values in concordant and discordant cell lines, Boxplots show the median and interquartile range; individual points represent cell lines.

Robustness of classification to threshold selection

The reliability of the classification was assessed by shifting the threshold values to the 40th and 60th percentiles of the distribution. Comparable results were observed: the proportion of discordant results between the LN(IC_{50}) and AUC metrics was 6.9% in both cases, which is comparable to the 8.6% figure obtained with median splitting (Table X).

Distribution of discordant cases relative to decision boundaries

A comparison of distances to the median revealed marked differences between the groups. For both

metrics (LN(IC_{50}) and AUC), discordant points were found to be significantly closer to the threshold ($p = 0.0001$), whilst concordant cell lines exhibited a wide range of values (Figure 7).

Statistical analysis using the Mann–Whitney test showed that discordant observations were located significantly closer to the classification boundary (Table XI).

Table X

Effect of threshold selection on discordance between LN(IC_{50}) and AUC classifications

Threshold (quantile)	Discordance (%)
0.4	6.9
0.5	8.6
0.6	6.9

Discordance is expressed as the percentage of cell lines with inconsistent classification at different quantile thresholds

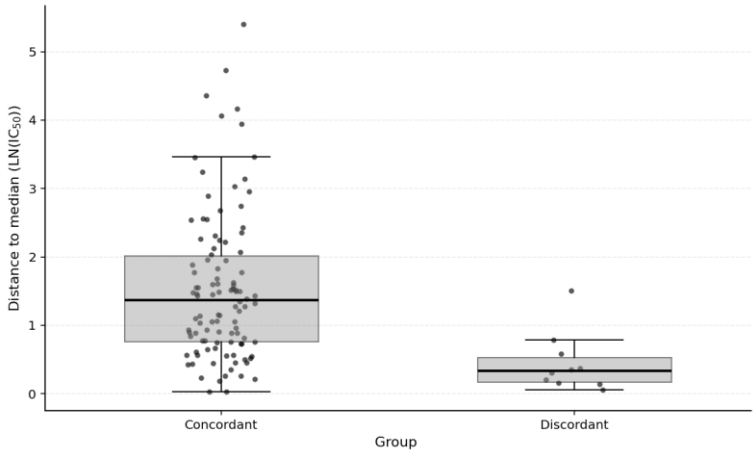

Figure 7.

Distance to the median LN(IC_{50}) in concordant and discordant cell lines. Distances were calculated as absolute deviations from the median. Boxplots show the median and interquartile range; individual points represent cell lines.

Table XI

Distance to the median for concordant and discordant cell lines

Group	Distance LN(IC_{50}), median (Q1–Q3)	Distance AUC, median (Q1–Q3)
Concordant	1.3683 (0.7570–2.0158)	0.0747 (0.0448–0.1028)
Discordant	0.3298 (0.1701–0.5277)	0.0041 (0.0018–0.0062)

Distances are expressed as median and interquartile range (Q1–Q3) for LN(IC_{50}) and AUC metric

Stability analysis showed that the proportion of discordant observations changed only modestly when alternative threshold values were applied. Median and alternative quantile partitions yielded comparable levels of discordance, suggesting that the observed findings were not solely dependent on the specific threshold selected for binary classification (4). Since the classification was based on a median split of two highly correlated measures, the concentration of discordant observations near the classification boundaries is an expected consequence of the chosen approach. However, it is not the mere fact of their localisation near the thresholds that is of practical interest, but rather a quantitative assessment of the extent of this vulnerable zone. Despite the very high correlation between LN(IC_{50}) and AUC ($\rho = 0.967$), a median split of these metrics resulted in 8.6% of cell lines being classified differently, indicating the potential sensitivity of binary stratification to the choice of metric. This finding is consistent with the statistical principle that dichotomisation of continuous variables should be interpreted with caution, particularly for observations close to the cut-off values (4, 24, 25).

Methodological strengths and robustness of the study

This study has several methodological advantages that ensure the reliability of its conclusions. All raw data were obtained from a single GDSC2 repository as part of a fixed pharmacological protocol, ensuring standardisation of the experimental conditions (26). The use of a 'single drug–multiple tumour types' study design allowed us to rule out the influence of inter-drug variability, whilst multilevel analysis enhanced the reliability of the interpretations. The analysis was not limited to statistical significance (p-values). The emphasis on the practical significance of differences helped to avoid false-positive conclusions, which are common when analysing large samples. Validation of the median estimates using bootstrap resampling and calculation of Cliff's delta contributed to the robustness and reliability of the analysis (27). The use of median-based classification provided a statistically neutral framework for

comparing the concordance of binary classifications derived from LN(IC_{50}) and AUC. The analysis of both continuous metrics alongside their binary classification allowed comparison of continuous and binary interpretations of drug response (28, 29). The parallel analysis of LN(IC_{50}) and AUC improved the transparency of phenotype classification and facilitated comparison of classification outcomes between the two metrics. A systematic interpretation of the results points to the robustness of the patterns identified in the models under investigation. The patterns identified highlight the need for a multi-parameter approach to the analysis of pharmacogenomic data. Such a comprehensive approach should take into account both continuous response metrics and the stability of binary classification in observations located near decision boundaries.

Limitations and future directions

Despite the consistency of the identified patterns, their analysis must take into account a number of objective limitations within which the study was conducted. The analysis focused on a single drug, camptothecin. We cannot extrapolate the identified inter-tumour differences to other classes of anticancer agents, as the pharmacological response is specific to each class of compounds and depends on the characteristics of its interaction with cellular targets (14). A priority for future research may be the verification of the observed patterns using other model agents with mechanisms of action different from those of camptothecin. In the present study, indicators of cytotoxic response were derived exclusively from *in vitro* models, which do not reflect the full complexity of the pharmacokinetic processes occurring *in vivo*. Further research could focus on integrating these findings with the results of analyses of more complex biological models. The factor of uneven group sizes under investigation may also affect the accuracy and reliability of intergroup comparisons, particularly in groups with small sample sizes (CESC, BLCA). In future studies, broadening the range of tumour groups selected and increasing their sample sizes could improve the accuracy of the results. The analysis in this study, despite the multi-stage nature of the independent methods, does not take into account the molecular-genetic characteristics of the cell lines. A deeper understanding of the mechanisms of intermolecular interaction could serve as a topic for further research. This study focused on the analysis of two indicators: LN(IC_{50}) and AUC. In the future, taking into account parameters such as maximum effect (E_{max}) and the Hill slope may provide a deeper and more detailed understanding of the identified discrepancies in phenotype classification (5, 30). Despite fixed pharmacological conditions, the proposed approach can be applied to other pharmacogenomic datasets and scaled up to larger pharmacogenomic datasets to verify the consistency of various cytotoxicity metrics (20, 21). Furthermore, the observed discordance in binary classification (8.6%), despite a high correlation between LN(IC_{50}) and AUC ($\rho = 0.97$), highlights the limitations of relying on a single response metric for phenotype classification and supports the joint interpretation of LN(IC_{50}) and AUC, particularly in borderline cases (20, 21).

Conclusions

Analysis of GDSC2 data demonstrated clear inter-tumour differences in sensitivity to camptothecin. BRCA cell lines showed reduced sensitivity compared with other tumour types, as reflected by higher LN(IC_{50}) and AUC values. Despite the high correlation between LN(IC_{50}) and AUC ($\rho = 0.97$), using these metrics for binary classification resulted in discordance in 8.6% of observations, rising to 11.8% in certain tumour groups. The results obtained show that cell lines situated close to the threshold values are most prone to a change in category when binary classification is used. The proposed strategy of joint analysis of LN(IC_{50}) and AUC can serve as a tool for more accurate interpretation of results and minimisation of the loss of biologically meaningful information. Accordingly, the results of the study highlight the importance of jointly analysing these metrics and of cautious interpretation of binary classification results, particularly for observations situated close to the threshold values. The proposed approach can be used to identify cases in which using different metrics yields different binary interpretations of sensitivity. This may contribute to a more informed interpretation of drug response.

Acknowledgement

Not applicable.

Funding

Not applicable.

Availability of data and materials

The data analysed in this study are publicly available in a precompiled snapshot of the GDSC2 dataset hosted on Kaggle (URL), downloaded on 13 August 2024.

Ethics approval and consent to participate

Not applicable. This study used publicly available, de-identified data from the GDSC2 database.

AI use disclosure

During the preparation of this manuscript, the authors used machine translation tools for partial translation from Russian and Portuguese into English,

followed by AI-assisted language editing to improve grammar, clarity, and readability. The authors reviewed and revised all outputs and take full responsibility for the final content of the work.

Conflict of interest

The authors declare no conflict of interest.

References

1. Garnett MJ, Edelman EJ, Heidorn SJ, Beckett P, Sanceau JY, Guertin IS, et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature. 2012;483(7391):570-575.
2. Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Arane MP, Sekine M, et al. A landscape of pharmacogenomic interactions in cancer. Cell. 2016;166(3):740-754.
3. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013;41(D1):D955-D961.
4. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ. 2006;332(7549):1080.
5. Fallahi-Sichani M, Honarnejad S, Heiser LM, Gray JW, Sorger PK. Metrics other than potency reveal systematic variation in responses to cancer drugs. Nat Chem Biol. 2013;9(11):708-714.
6. Hafner M, Niepel M, Chung M, Sorger PK. Growth rate inhibition metrics correct for confounders in measuring sensitivity to cancer drugs. Nat Methods. 2016;13(6):521-527.
7. McGill R, Tukey JW, Larsen WA. Variations of box plots. Am Stat. 1978;32(1):12-16.
8. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. J Am Stat Assoc. 1952;47(260):583-621.
9. Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. Ann Math Stat. 1947;18(1):50-60.
10. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat. 1979;6(2):65-70.
11. Cliff N. Dominance statistics: ordinal analyses to answer ordinal questions. Psychol Bull. 1993;114(3):494-509.
12. Spearman C. The proof and measurement of association between two things. Am J Psychol. 1904;15(1):72-101.
13. Efron B, Tibshirani RJ. An introduction to the bootstrap. 1st ed. New York: Chapman & Hall; 1993.
14. Pommier Y. Topoisomerase I inhibitors: camptothecins and beyond. Nat Rev Cancer. 2006;6(10):789-802.
15. Yan F, Robert M, Li Y. Statistical methods and common problems in medical or biomedical science research. Int J Physiol Pathophysiol Pharmacol. 2017;9(5):157-163.
16. Desai SD, Li TK, Rodriguez-Bauman A, Rubin EH, Liu LF. Ubiquitin/26S proteasome-mediated degradation of topoisomerase I as a determinant of camptothecin sensitivity. Cancer Res. 2001;61(15):5926-5932.
17. Tesauro C, Simonsen AK, Andersen MB, Petersen KW, Kristoffersen EL, Algreen L, et al. Topoisomerase I activity and sensitivity to camptothecin in breast cancer-derived cells. BMC Cancer. 2019;19(1):1158.
18. Basu A, Bodycombe NE, Cheah JH, Price EV, Liu K, Schaefer GI, et al. An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. Cell. 2013;154(5):1151-1161.
19. Seashore-Ludlow B, Rees MG, Cheah JH, Cokol M, Price EV, Coletti ME, et al. Harnessing connectivity in a large-scale small-molecule sensitivity dataset. Cancer Discov. 2015;5(11):1210-1223.
20. Sharifi-Noghabi H, Jahangiri-Tazehkand S, Smirnov P, Hon C, Mammoliti A, Nair SK, et al. Drug sensitivity prediction from cell line-based pharmacogenomics data: a review. Brief Bioinform. 2021;22(6):bbab098.
21. Hafner M, Niepel M, Sorger PK. Alternative drug sensitivity metrics improve preclinical cancer pharmacogenomics. Nat Biotechnol. 2017;35(6):500-502.
22. Ghandi M, Huang FW, Jané-Valbuena J, Kryukov GV, Lo CC, McDonald ER, et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. Nature. 2019;569(7757):503-508.
23. Bland JM, Altman DG. Statistics Notes: Bootstrap resampling methods. BMJ. 2015;350:h2622.
24. Busch EL. Cut points and contexts. Cancer. 2021;127(23):4314-4317.
25. Polley MYC, Dignam JJ. Statistical considerations in the evaluation of continuous biomarkers. J Nucl Med. 2021;62(5):605-611.
26. Hatzis C, Bedard PL, Birkbak NJ, Beck AH, Aerts HJ, Stem DF, et al. Enhancing reproducibility in cancer drug screening: how do we move forward? Cancer Res. 2014;74(15):4016-4023.
27. Ploenzke M, Irizarry R. Reassessing pharmacogenomic cell sensitivity with multilevel statistical models. Biostatistics. 2023;24(4):901-921.
28. Mpindi JP, Yadav B, Östling P, Gautam P, Malani D, Murumägi A, et al. Consistency in drug response profiling. Nature. 2016;540(7631):E5-E6.
29. Safikhani Z, Smirnov P, Freeman M, El-Hachem N, She A, Goldenberg A, et al. Revisiting inconsistency in pharmacogenomic studies. F1000Res. 2016;5:2333.
30. Brooks EA, Galarza S, Gencoglu MF, Cornelison RC, Munson JM, Peyton SR. Applicability of drug response metrics for cancer studies using biomaterials. Philos Trans R Soc Lond B Biol Sci. 2019;374(1779):20180226.