Published in final edited form as:

Nat Biotechnol. 2017 June 07; 35(6): 500–502. doi:10.1038/nbt.3882.

Alternative drug sensitivity metrics improve preclinical cancer pharmacogenomics

Marc Hafner, Mario Niepel, and Peter K. Sorger*

HMS LINCS Center, Laboratory of Systems Pharmacology, Department of Systems Biology, Harvard Medical School, Boston, MA, USA

Measuring drug response in cancer cell lines is essential for studying mechanisms of drug action and identifying genetic variants associated with sensitivity and resistance (preclinical pharmacogenomics). The conventional approach involves computing relative viability based on the ratio of cells in control and drug-treated cultures at the end of a fixed time period (Fig. 1a and Supplemental Fig. 1a). We have recently shown that relative viability and response metrics derived from it, such as IC_{50} and area under the dose-response curve (AUC), are confounded by variation in cell proliferation rates and assay duration¹, potentially explaining discrepancies among large-scale drug response datasets^{2,3}. Here we show that use of relative viability as measure of drug response results in false-positive and -negative pharmacogenomic associations. An alternative approach, based on computing normalized growth rate (GR) inhibition minimizes these artefactual associations by scoring drug sensitivity on a per-division basis¹ (Fig. 1a). GR_{50} , a measure of potency, is the concentration of drug that reduces cell proliferation rate by one-half; GR_{max} , a measure of efficacy, is the maximum effect of a drug at the highest tested concentration; and GR_{AOC} combines these in an integrated “area over the curve” value. The signs of GR values and GR_{max} relate directly to response phenotype: positive for partial growth inhibition, zero for complete cytostasis, and negative for cell death. By contrast, E_{max} values do not have this property as they are strongly confounded by proliferation rates^{1,4}. Collecting GR values requires only modest changes in experimental approach^{5,6} and calculations can be performed online (GRcalculator.org).

A recently released dataset from the Genentech Cell Line Screening Initiative (gCSI)⁷ is unique among large-scale drug response datasets in reporting cell division times, which are required for GR calculations. Median cell line division times varied ~2-fold across tissue types and up to 3-fold within cell lines of a given type (note that division times also vary with culture conditions^{1,2}): lymphoid cancer lines had the shortest median division times at 29 h with an interquartile range (IQR) of 15 h, whereas breast cancer lines had the longest division times at 53 h (IQR=16 h; Fig. 1b). All other things being equal, slower division rates increase IC_{50} and E_{max} values for simple arithmetic reasons (Fig. 1a).¹ In the gCSI dataset, both IC_{50} and E_{max} were in fact substantially higher for the slowest dividing cell types relative to the fastest (Spearman's $\rho>0.61$, $p<0.005$). For drugs such as paclitaxel,

*Corresponding author: Peter K. Sorger, 200 Longwood Avenue, Boston, MA 02115, peter_sorger@hms.harvard.edu, phone (617) 432-6902, fax (617) 432-6990.

Supplemental Data: Calculated GR values and GR metrics for the gCSI and the CTRP datasets.

docetaxel, and gemcitabine, drug sensitivity also correlated with cell division times ($p > 0.35$, $p < 5.6 \times 10^{-13}$ for IC_{50} , $p > 0.4$, $p < 1.4 \times 10^{-16}$ for E_{max}). Based on IC_{50} values, bone cancer cells were reportedly 60-fold more drug sensitive on average than breast cancer cells even though the drugs in the gCSI dataset are primarily used to treat the latter disease. By contrast, we found that median GR_{50} and GR_{max} values are not significantly correlated across fast- and slow-growing tumor cell types ($p > 0.20$). Thus, the correlation between tumor type and median IC_{50} or E_{max} values arises from systematic variation in division time.

Conventional drug sensitivity metrics are similarly confounded by genetic mutations (natural and engineered) that alter proliferation rates^{8,9} (Supplemental Fig. 1b-d) resulting in both bona fide and confounded pharmacogenomic associations. Overall, only 40% of 122 statistically significant associations ($FDR < 0.15$) between IC_{50} values and genomic alterations (gene mutations, deletions, or amplifications) in gCSI data were significant when using GR_{50} values (Fig. 1c). Conversely, about half of associations scored as significant by GR_{50} were missed using IC_{50} . An example of a false-negative association is provided by PTEN loss-of-function (LOF) mutations, which are known to mediate lapatinib resistance in breast cancer cells¹⁰ but were not associated with differences in IC_{50} values ($p = 0.12$, rank-sum test) simply because PTEN mutant cell lines also had a median division time that was 22.5 h shorter than PTEN wild-type cell lines ($p = 0.01$), consistent with the known role of PTEN as a tumor suppressor. By contrast, the association between PTEN LOF and lapatinib resistance was significant when GR_{50} values were used ($p = 0.003$; Fig. 1d). A false-positive example is provided by deletion of CDC73 (a component of an RNA Polymerase II complex¹¹), which was associated in ovarian cancer lines with a 1000-fold increase in median IC_{50} for docetaxel ($p = 0.002$; Fig. 1e). CDC73-deleted cells grew more slowly than CDC73 expressers (57 vs. 35 h doubling time; $p = 0.03$), and when we corrected for this difference using GR_{50} values, the association between CDC73 deletion and docetaxel sensitivity disappeared ($p = 0.13$). We conclude that differences in cell division rate underlie false-positive and -negative associations between genomic features and conventional drug response metrics.

Efficacy is rarely explored as a metric of drug response, even though incomplete cytostasis and fractional cell killing result in persistence of viable cells¹² that likely contributes to incomplete tumor regression and residual disease¹³. Across the gCSI dataset both potency and efficacy exhibited substantial variation, and even highly potent drugs could be inefficient at killing (Fig. 2a, Supplemental Fig. 2). For example, the mTOR inhibitor rapamycin was potent ($GR_{50} < 0.1 \mu M$) in 34% of cell lines but ineffective in blocking growth in these lines ($GR_{max} = 0.34 \pm 0.23$, median and IQR). The converse was also true: the HDAC inhibitor vorinostat was highly cytotoxic in most lines ($GR_{max} = -0.52 \pm 0.48$) but had poor potency (median $GR_{50} = 1.53 \mu M$). The microtubule inhibitor docetaxel was highly potent (median $GR_{50} = 8 \text{ nM}$) but cytotoxic in only half of cell lines tested ($GR_{max} = -0.04 \pm 0.53$). For doxorubicin, which is commonly used in hematopoietic cancers¹⁴, GR_{max} values were significantly lower in blood cell lines than in cell lines from other tissues (median of -0.84 vs. -0.49 , $p = 1.6 \times 10^{-6}$) but GR_{50} values were similar ($p = 0.14$).

We identified 99 statistically significant associations ($FDR < 0.15$) between genomic variants and GR_{max} values and the same number for GR_{50} , but only 18 of those variants were

significantly associated with both GR_{max} and GR_{50} (Fig. 2b). For example, in most ovarian cell lines deleted for the anti-apoptotic BCL2 gene, GR_{max} values for docetaxel were negative, which indicates cell death, whereas GR_{max} values in BCL2 wild-type lines were positive, which indicates partial cytostasis (median of -0.18 vs. 0.23 ; $p=0.007$; Fig. 2c). By contrast, GR_{50} values were indistinguishable ($p=0.23$). The sensitivity of BCL2-deleted ovarian cell lines to docetaxel as measured by GR_{max} is consistent with an observed increase in apoptosis when the BCL2 inhibitor ABT737 is combined with docetaxel¹⁵.

To study the relationship between GR_{50} and GR_{max} in a systematic manner, we analyzed the CTRP dataset¹⁶, which comprises 545 drugs but lacks data on cell proliferation rates. We inferred division times for 121 cell lines using gemcitabine response as a fiducial (see methods). In place of GR_{max} we used an interpolated value $GR_{2\mu M}$ based on fitted dose-response curves. Shannon's entropy (H) was then used to measure the information content in distributions of GR_{50} and $GR_{2\mu M}$ values for each drug. We observed a wide range of entropies with some drugs varying primarily in GR_{50} , some in $GR_{2\mu M}$, and some in both (Supplemental Fig. 3a). The mutual information between GR_{50} and $GR_{2\mu M}$ was low, showing that the two metrics of drug response convey non-overlapping information. AUC is often used as a summary statistic for a drug, but we found that the entropy of GR_{AOC} , the GR equivalent of AUC , is lower than the joint entropy of GR_{50} and $GR_{2\mu M}$ (Supplemental Fig. 3b, c). This makes sense because the same GR_{AOC} value can be associated with a potent drug having poor efficacy or a less potent drug having good efficacy (Supplemental Fig. 3d).

We conclude that GR metrics have the potential to improve pre-clinical pharmacology in three ways. First, use of IC_{50} values based on relative viability in pharmacogenomic studies results in multiple false- positive and -negative associations simply because cell division rates vary in a systematic manner with tumor type and genotype. Artefactual associations arise, regardless of experimental methodology or type of disease, whenever cells divide over the course of a drug-response assay and the phenotype of interest is directly or indirectly affected by growth. GR metrics correct for this and also for arbitrary differences in protocol that affect cell division number (e.g. assay duration, media conditions, plating density). Second, because GR metrics reduce the impact of confounders, cross-study reproducibility is improved: for example, the correlation between gCSI and CTRP datasets is significantly increased using GR_{AOC} rather than AUC as a response metric ($p=1.3\times 10^{-3}$, t-test; Supplemental Fig. 4). Third, efficacy as measured by GR_{max} and potency as measured by GR_{50} differ at a biological level, carry complementary information (low mutual information), and are associated with largely non-overlapping genetic alterations. In principle, variation in potency and efficacy can be captured by integrating across dose-response curves (GR_{AOC}), but we find that information content is maximized if GR_{50} and GR_{max} are considered independently.

Because the ultimate purpose of antineoplastic drugs is to kill cancer cells¹³ and high potency is no guarantee of good efficacy, we propose that the best drugs and most important pharmacogenomic associations are not those associated with low IC_{50} values but rather those that result in the most negative GR value at clinically relevant drug concentrations (e.g. C_{max}). Relating *in vitro* measures of drug sensitivity to *in vivo* responses remains

challenging¹⁷, but for this to have any chance of success it is essential that *in vitro* data are as informative and reproducible as possible.

Online Methods

GR values and metrics were calculated from the gCSI dataset based on the equation:

$$GR(c) = 2^{1 + \frac{\log_2(\bar{x}(c))}{T/T_d}} - 1,$$

where $\bar{x}(c) = x(c)/x_{ctrl}$ is the relative viability, T_d is the reported division time, and $T=72$ h is the assay length (see Supplemental Fig. 1 and Hafner *et al.*¹). Note that this equation can be applied to calculate GR values from any conventional dataset for which the division time of the cell lines is known. Associations are based on the list of oncogenes and tumor suppressor genes published in Vogelstein *et al.*¹⁸ and the genomic alterations published on the CCLE portal (binary call data). For enrichment analyses and correlations by tissue type, only tissues having at least four cell lines in the dataset were considered. Enrichment significance was evaluated using Wilcoxon's rank-sum test; FDR values were calculated by applying the Benjamini and Hochberg procedure independently for each combination of tissue, drug, and metric. For enrichment analyses, only comparisons with at least four cell lines per group were considered. GR values and metrics are reported in the supplemental data and accessible via the interactive website GRbrowser.org.

To calculate GR values for the CTRP data¹⁶, we inferred the cell count at the time of treatment in the CTRP data relative to the untreated control at the end of the treatment,

$\bar{x}_0^{CTRP} = x_0/x_{ctrl}$, by matching a reference GR value, GR_{ref}^{gCSI} , from a drug treatment in the gCSI data⁷ to a reference relative cell count value, E_{ref}^{CTRP} , from the same treatment in the CTRP data with the formula:

$$\bar{x}_0^{CTRP} = E_{ref}^{CTRP} \wedge \left(1 / \left(1 - \log_2 \left(GR_{ref}^{gCSI} + 1 \right) \right) \right)$$

Gemcitabine is a particularly good candidate as a reference because it induces a cytostatic response in most cell lines with a relatively flat plateau at high concentrations, which helps with curve-fitting and small differences in potency between the datasets. We defined the

GR_{ref}^{gCSI} and E_{ref}^{CTRP} values as the interquartile mean across all measurements made at concentrations ranging from 0.1 μ M to 32 μ M in the respective datasets. Note that we discarded measurements in the CTRP data obtained at drug concentrations higher than 32 μ M because these were noisier. To avoid artefacts, we only considered cell lines for which (1) GR_{ref}^{gCSI} value for gemcitabine in was below 0.7; (2) E_{ref}^{CTRP} value for gemcitabine was below 0.85; (3) gemcitabine response in the CTRP data could be fitted to a sigmoidal curve with $r^2 > 0.5$; (4) the difference between GEC_{50} value in the gCSI data and EC_{50} value in the CTRP data was less than 1.5 orders of magnitude; and (5) the variance of the GR values of the gCSI data and relative cell count values of the CTRP data in the range 0.1 μ M to 32 μ M

was less than 0.01. Among the 332 cell lines common to the two datasets, 44% had gemcitabine data of sufficient quality to provide reliable inference of the number of cells at time of treatment and thus enable calculation of GR values. Given the inferred $\bar{x}_0^{\text{CTRP}}$ and the relative viability $\bar{x}(c)$ reported in the CTRP data, all GR values and metrics were calculated using the equation:

$$GR(c) = 2^{\frac{\log_2(\bar{x}(c)/\bar{x}_0^{\text{CTRP}})}{\log_2(1/\bar{x}_0^{\text{CTRP}})}} - 1$$

For the CTRP data, we used $GR_{2\mu\text{M}}$ (GR value at 2 μM) in place of GR_{max} because the highest concentrations used in the CTRP experiments are not biologically relevant and, as noted above, the measurements at high concentrations were noisier.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Author contributions: MH, MN, and PKS conceived this study and wrote the paper. MH performed the analyses. This work was supported by grant U54-HL127365. We thank J. W. Gray, R. Bourgon and colleagues at Genentech, and S. Schreiber for helpful discussions.

Bibliography

1. Hafner M, Niepel M, Chung M, Sorger PK. Nat Methods. 2016; 13:521–7. [PubMed: 27135972]
2. Haibe-Kains B, et al. Nature. 2013; 504:389–93. [PubMed: 24284626]
3. CCLE Consortium & GDSC Consortium. Nature. 2015; 528:84–7. [PubMed: 26570998]
4. Harris LA, et al. Nat Methods. 2016; 13:497–500. [PubMed: 27135974]
5. Niepel M, Hafner M, Chung M, Sorger PK. Curr Protoc Chem Biol. 2017; 9
6. Hafner M, Niepel M, Subramanian K, Sorger PK. Curr Protoc Chem Biol. 2017; 9
7. Haverty PM, et al. Nature. 2016; 533:333–7. [PubMed: 27193678]
8. Im HK, et al. PLoS Genet. 2012; 8:e1002525. [PubMed: 22346769]
9. Choy E, et al. PLoS Genet. 2008; 4:e1000287. [PubMed: 19043577]
10. Eichhorn PJ, et al. Cancer Res. 2008; 68:9221–30. [PubMed: 19010894]
11. Rozenblatt-Rosen O, et al. Mol Cell Biol. 2005; 25:612–20. [PubMed: 15632063]
12. Fallahi-Sichani M, et al. Nat Chem Biol. 2013; 9:708–14. [PubMed: 24013279]
13. Letai A. Clin Cancer Res. 2015; 21:5015–20. [PubMed: 26567360]
14. Fisher R, et al. N Engl J Med. 1993; 328:1002–6. [PubMed: 7680764]
15. Oakes SR, et al. Proc Natl Acad Sci U S A. 2012; 109:2766–71. [PubMed: 21768359]
16. Seashore-Ludlow B, et al. Cancer Discov. 2015; 5:1210–23. [PubMed: 26482930]
17. Johnson JJ, et al. Br J Cancer. 2001; 84:1424–31. [PubMed: 11355958]
18. Vogelstein B, et al. Science. 2013; 339:1546–58. [PubMed: 23539594]

Figure 1. Division rate confounds IC_{50} values and pharmacogenomic associations

(a) Theoretical dose-response curves by relative cell count (left) and GR value (right) for cell lines with slow ($T_d=70$ h; solid lines) or fast ($T_d=35$ h; dashed lines) division times treated with a drug that partially inhibits division (blue) or causes cell death (green). Slow- and fast-growing lines are assumed to respond identically at a mechanistic level. IC_{50} and GR_{50} are marked on the x-axis; $E_{\max}$ and $GR_{\max}$ on the y-axis. Note that dashed and solid lines overlap in the GR value panel (see also Supplemental Fig. 1a). (b) Distribution of division time for cell lines by tissue of origin in the gCSI dataset. (c) IC_{50} and GR_{50} enrichment by

genomic alteration. Numbers represent significant ($FDR < 0.15$) associations for IC_{50} alone, GR_{50} alone, or IC_{50} plus GR_{50} . (d-e) Distribution of IC_{50} and GR_{50} values for (d) lapatinib in breast lines based on PTEN status and (e) docetaxel in ovarian lines based on CDC73 status. Rank-sum p-values are reported. GR_{50} and IC_{50} values are capped at 31 μM .

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2. Variation in efficacy and genomic enrichment based on GR_{max} values in the gCSI dataset

(a) Distribution of GR_{max} and GR_{50} values for each drug across all cell lines. GR_{50} values are capped at 31 μM . (b) GR_{50} and GR_{max} enrichment by genomic alteration for each drug and tissue. Numbers represent significant ($FDR < 0.15$) associations for GR_{max} alone, GR_{50} alone, or GR_{max} plus GR_{50} . (c) Distribution of GR_{max} values for docetaxel in ovarian cancer lines based on BCL2 deletion status. Rank-sum p-value is reported.