

SCIENTIFIC REPORTS

OPEN

Estimating Potency in High-Throughput Screening Experiments by Maximizing the Rate of Change in Weighted Shannon Entropy

Received: 04 April 2016

Accepted: 26 May 2016

Published: 15 June 2016

Keith R. Shockley

High-throughput *in vitro* screening experiments can be used to generate concentration-response data for large chemical libraries. It is often desirable to estimate the concentration needed to achieve a particular effect, or potency, for each chemical tested in an assay. Potency estimates can be used to directly compare chemical profiles and prioritize compounds for confirmation studies, or employed as input data for prediction modeling and association mapping. The concentration for half-maximal activity derived from the Hill equation model (i.e., AC_{50}) is the most common potency measure applied in pharmacological research and toxicity testing. However, the AC_{50} parameter is subject to large uncertainty for many concentration-response relationships. In this study we introduce a new measure of potency based on a weighted Shannon entropy measure termed the weighted entropy score (*WES*). Our potency estimator (Point of Departure, POD_{WES}) is defined as the concentration producing the maximum rate of change in weighted entropy along a concentration-response profile. This approach provides a new tool for potency estimation that does not depend on the assumption of monotonicity or any other pre-specified concentration-response relationship. POD_{WES} estimates potency with greater precision and less bias compared to the conventional AC_{50} assessed across a range of simulated conditions.

Quantitative high-throughput screening (qHTS) assays¹ return thousands of concentration-response profiles for large chemical libraries and are currently driving major advancements in drug discovery² and toxicity testing³. For example, more than 10,000 substances are now being tested in 15-point concentration-response format in phase II of the Tox21 collaboration, involving the U.S. Environmental Protection Agency (EPA), the U.S. Food and Drug Administration (FDA), the National Institutes of Health (NIH) National Center for Advancing Translational Sciences (NCATS) and the National Institute for Environmental Health Sciences (NIEHS)/National Toxicology Program (NTP)⁴. Response profiles can be summarized by a measure of average activity across tested concentrations, such as the area under the curve (*AUC*) of concentration-response curves⁵, a weighted version of *AUC*⁶, or a weighted entropy score (*WES*)⁷. While these measures are useful for ranking compounds, it is often desirable to estimate the concentration at which a chemical induces a particular effect level using automated data analysis processes. Such potency measures can be applied for rapid identification of pharmacologically active hits or toxicological assessment, or used as input data for prediction modeling⁸ or association mapping⁵.

The most common approach used to approximate chemical potency in chemical genomics and large-scale toxicity testing is the AC_{50} parameter in the Hill Equation model⁹. The AC_{50} parameter estimates the concentration at which a chemical produces the half-maximal response along a sigmoidal curve¹⁰. Incorporating domain knowledge into the curve fitting process can improve agreement between AC_{50} estimates for sigmoidal curves¹¹. However, it is not possible to know the underlying shape of the concentration-response relationship before conducting an experiment¹² and complex response patterns may reflect real biological responses¹³. Furthermore, linearizing assumptions can render AC_{50} parameter estimation from the Hill model very unreliable, even with increased sample sizes^{10,14}. Applying individualized curve fitting procedures can be useful for characterizing screening results. However, in the high-throughput setting manual scrutiny can be restrictively laborious and

Biostatistics and Computational Biology Branch, The National Institute of Environmental Health Sciences, National Institutes of Health, 111 T. W. Alexander Drive, Research Triangle Park, NC 27709, USA. Correspondence and requests for materials should be addressed to K.R.S. (email: shockleykr@niehs.nih.gov)

result in extensive data censoring. Also, while outlier removal and parameter constraints may reduce curve fit error, these procedures do not necessarily increase the repeatability of nonlinear parameter estimation. It is not unusual for AC_{50} estimates to be accompanied by large standard errors even when one or both asymptotes can be defined¹⁰.

A point of departure (POD) represents a concentration derived from observed concentration-response data that is associated with a defined effect. *In vitro* POD estimates have been calculated based on linear interpolation between the two concentrations that lie on either side of the assay detection threshold⁶ or establishing a baseline noiseband using the first two tested concentrations¹⁵. Other POD metrics include an estimate of the concentration producing a predetermined level of an adverse response (i.e., the benchmark dose or BMD) and the highest tested concentration for which there is no observed adverse effect (i.e., the no-observed-adverse-effect-level or NOAEL)^{16,17}. With true experimental replicates, BMD modeling or NOAEL determinations could serve as POD estimates describing the concentration at which the assay response begins to deviate from baseline response levels. Unlike the NOAEL approach, the BMD procedure uses mathematical modeling to make use of the entire observed concentration-response profile. Unfortunately, in qHTS studies there is usually very little, if any, replication at each tested concentration and it is often not appropriate to combine data across different experimental runs because conditions can change substantially between trials^{4,10}.

We propose a nonparametric approach based on information theory to improve the precision of compound potency estimation in qHTS studies. Information theoretic concepts were originally developed for communication technology¹⁸, but these approaches have recently been used to summarize patterns in gene expression microarray data^{19,20}, find differential methylation sites²¹ and rank chemicals in qHTS experiments⁷. Shannon entropy (H) describes the average information content in a probability distribution²², and can be used to describe the extent and uniformity of response in a concentration-response profile. Here, H is computed from the probability distribution obtained from the observed responses and naturally accommodates any concentration-response pattern, not just monotonic trends such as the sigmoidal shape of the Hill equation model.

We define compound potency as the concentration producing the maximal rate of change in entropy. This potency is calculated by finding the maximum first derivative of the entropy measure across the concentration range. However, Shannon entropy does not take into account the uncertainty in response measurements when responses are within the noise region, i.e., measurements that are less than the assay detection limit. We therefore employ a weighted version of Shannon entropy (or WES)⁷. WES weights responses found within the noise region so that profiles with larger WES scores have greater probability mass (i.e., greater average activity) in the detectable region of the assay. Accordingly, the point of departure is found at the concentration where the rate of change in weighted entropy is maximized along the tested concentration range. This new potency estimator is termed POD_{WES} . Unlike the AC_{50} value, POD_{WES} does not rely on the shape of the profile far removed from the point of departure. Observed concentration-response profiles that lie entirely within the assay noise region are assigned the outcome “undefined”. Profiles which have detectable responses and for which the maximum rate of change in weighted entropy is located at the lowest observed concentration C_1 , where POD_{WES} must be less than C_1 but cannot be estimated from the given data, are assigned the outcome “less than C_1 ”.

Results

Computing POD_{WES} for illustrative profiles. Figure 1 summarizes the workflow used to calculate POD_{WES} . To begin, WES and its derivatives are calculated at each tested concentration level. Chemicals with larger WES scores have greater average relative responses across concentrations⁷. If the maximum observed response is less than the assay detection limit, POD_{WES} is “undefined”, since a detectable response may have occurred if a larger range of test concentrations had been used. If at least one measured response is detectable, a search for a maximal rate of change in WES is conducted within the observed concentration-response space. If a global $\left| \frac{dWES}{dC} \right|_{\max}$ extremum is located, POD_{WES} is estimated. However, if POD_{WES} cannot be found, the concentration-response data is extrapolated outside of the observed concentration-response region using finite difference calculus. After extrapolating new responses, WES and its derivatives are recalculated and another search for POD_{WES} is conducted. If POD_{WES} still cannot be quantitatively determined, but $\left| \frac{dWES}{dC} \right|_{\max}$ is located at the lowest concentration in the extrapolated profile, POD_{WES} must be less than the lowest tested concentration (see Supplementary Information).

Figure 2 depicts hypothetical sigmoidal response profiles for three chemicals. Each chemical follows equation (1) in the Methods with no $ERROR$. The baseline response $R0$ is set to 0% of positive control, the maximal response $[RMAX]$ is set to 100% of positive control, the h parameter is set to 1 and the AC_{50} is set to 0.001, 0.1, and 10 μM , for *Chemical-1*, *Chemical-2* and *Chemical-3*, respectively. This figure shows the normalized responses (row 1), WES computed at each concentration level (row 2), the first derivative of WES at each concentration level (row 3) and the second derivative of WES at each concentration level (row 4). The concentration at which the first derivative of WES is maximized is indicated by an open square.

Chemical-1 is the most potent of the three chemicals shown in Fig. 2, where only the upper asymptote is well defined. This chemical has a “true” AC_{50} value equal to $1.00 \times 10^{-3} \mu\text{M}$, which corresponds to an POD_{WES} of $4.19 \times 10^{-4} \mu\text{M}$. *Chemical-2* has two clearly defined asymptotes with an AC_{50} value of 0.1 μM and a calculated POD_{WES} of 0.07 μM . One data point, indicated by an open triangle, was extrapolated in order to find POD_{WES} for *Chemical-3*, which had an AC_{50} value of 10 μM and a calculated POD_{WES} of 3.73 μM . In this case, a single extrapolated data point was used in order to calculate the deviation of the estimate of $\frac{d^2WES}{dC^2}$ from zero within the pre-specified tolerance level (see Supplementary Information for more explanation of the computations). Notice that the value of WES at the k th concentration level becomes smaller as the AC_{50} of a profile increases, but the potency

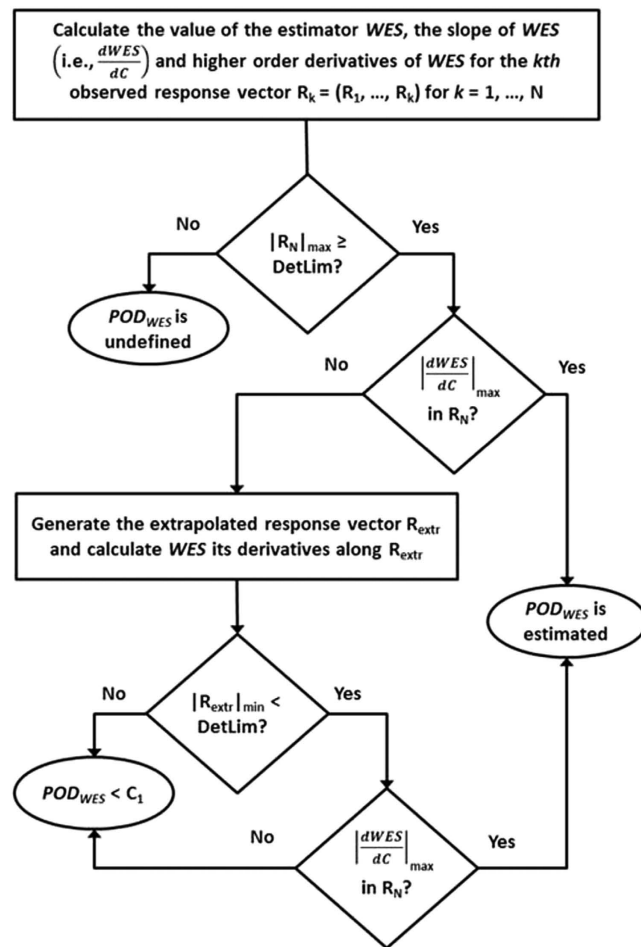


Figure 1. Overview of the potency calculation procedure. *WES* is the weighted entropy score, \mathbf{R}_k is the vector of observed response values (R_1, \dots, R_k), C_N is the vector of N tested concentrations with \mathbf{R}_N observed responses, $\frac{dWES}{dC}$ represents the first derivative of *WES* with respect to concentration, and \mathbf{R}_{extr} is the vector of response values obtained after data extrapolation.

measure POD_{WES} is located at the concentration for which the the rate of change in *WES* is increasing most rapidly. Fig. S1 shows additional examples of POD_{WES} calculated from curves generated with the “gain-loss” model given in equation (2) in the Methods.

Evaluating the proposed approach using simulated data. To explore precision and bias of the potency estimates derived from sigmoidal models, we generated 15-point concentration-response profiles from equation (1) in the Methods with $R_0 = 0\%$ and $h = 1$ for profiles having [1] only an upper asymptote ($AC_{50} = 0.001 \mu\text{M}$), [2] both asymptotes well defined ($AC_{50} = 0.1 \mu\text{M}$) and [3] only a lower asymptote ($AC_{50} = 10 \mu\text{M}$). In the simulations, $|R_{MAX}|$ values were selected as weak ($|R_{MAX}| = 25\%$), moderate ($|R_{MAX}| = 50\%$), and strong ($|R_{MAX}| = 100\%$) activity. A total of 10,000 profiles were generated for each of these nine combinations of AC_{50} and $|R_{MAX}|$ and the residual errors were modeled as $ERROR \sim N(\mu = 0, \sigma_i^2)$ where $\sigma_i = 5\%$ or $\sigma_i = 10\%$. In Table 1, the precision of potency estimation differed markedly between the estimators for the lower error of $\sigma_i = 5\%$. POD_{WES} estimates were generally more repeatable with confidence interval widths (CIWs) ranging from 1.03 to 1.53 orders of magnitude (OM). AC_{50} at the same error level ranged from 0.27 to 13.80 OM. The precision of POD_{WES} at $\sigma_i = 10\%$ was comparable to the levels achieved at $\sigma_i = 5\%$ for curves simulated under conditions in which “true” maximum response is greater than the detection limit of the assay. By contrast, precision of the AC_{50} estimator was noticeably lower for $\sigma_i = 10\%$ compared to $\sigma_i = 5\%$.

As shown in Table 1, for $\sigma_i = 5\%$, the bias in POD_{WES} estimates was less than about 2.0-fold on the natural scale, ranging from $\log_{10}\text{Bias} = -0.03$ (<1.1 -fold less than expectation) to $\log_{10}\text{Bias} = -0.27$ (<1.9 -fold less than expectation). For the same data sets, the bias in AC_{50} estimates ranged from $\log_{10}\text{Bias} = -0.0002$ (<1.1 -fold less than expectation) to $\log_{10}\text{Bias} = 1.38$ (24.0-fold greater than expected). The estimation bias of POD_{WES} at $\sigma_i = 10\%$ was similar to the values found at $\sigma_i = 5\%$. By contrast, the estimation bias of AC_{50} was about 10-fold greater for $\sigma_i = 10\%$ compared to $\sigma_i = 5\%$ in some instances.

In addition to investigating the precision and bias of potency estimators based on the sigmoidal Hill model, we also investigated the precision and bias of estimators using the “gain-loss” model from equation (2) in the

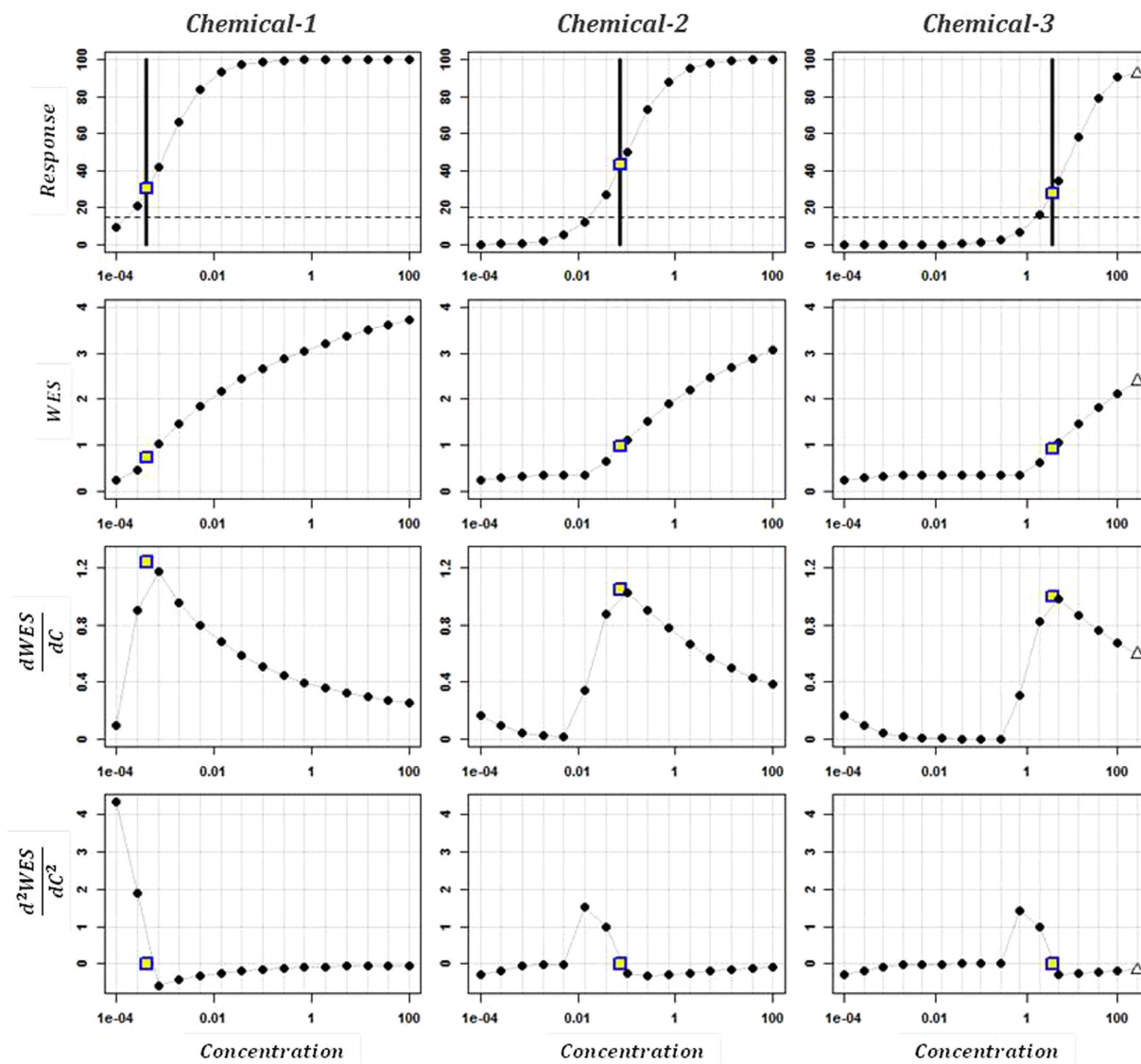


Figure 2. Illustrative example of the proposed approach to calculate potency for three 15-point concentration-response curves generated from the Hill equation model in equation (1) in the Methods with $R_{MAX} = 100\%$ of the positive control, $R_0 = 0$, the h parameter = 1 and $AC_{50} = 0.001, 0.1, \text{ and } 10$ for columns 1, 2, and 3, respectively. In the first row, responses are connected by gray lines, where solid circles represent observed responses, an open triangle shows an extrapolated response and dashed lines indicate a detection limit of 15%. Open squares indicate the estimated potency (POD_{WES}). Black vertical bars help to locate the estimated potency on the “Response” graphs. The first row shows the concentration-response, the second row indicates the values of WES at each response (i.e., WES at concentration k is computed by considering only the first 1, ..., k concentration levels), the third row shows the rate of change in WES at each concentration level, and the fourth row indicates the second derivative of WES at each concentration level.

Methods. As shown in Table 2, the precision of POD_{WES} was less than 1.5 OM at $\sigma_i = 5\%$ and $\sigma_i = 10\%$. By contrast, the AC_{50} measure could be extremely imprecise for these curve shapes, even reaching 19.78 OM in one case. Similar to the evaluation of Hill model curves, the \log_{10} Bias of POD_{WES} did not exceed -0.42 (<2.7 -fold less than expectation). The bias of the AC_{50} metric was often greater than 2.0 OM (or 100-fold).

Example compound potency estimates across experimental runs. Previously, the *in vitro* BG1 estrogen receptor alpha ($ER\alpha$) assay from phase II of Tox21 was used to screen for agonist activity in an ER reporter gene cell line with an endogenous full length $ER\alpha$. Approximately 10,000 compounds were assayed in three different experimental runs and activity measurements for 15-point concentration-response curves were obtained as luciferase activity readings from the BG1 $ER\alpha$ cell line²³. This data was normalized to 100% of the activity of estrone positive control compounds. Weighted entropy scores (WES) and POD_{WES} values were calculated as described here. Ranking profiles based on WES is not based on any pre-specified concentration response model form or direction of response⁷.

| Data Set Parameters | | | Estimator | |
|---------------------------------|----------------|------------------|----------------|--------------|
| True $ RMAX $ | True AC_{50} | True POD_{WES} | AC_{50} | POD_{WES} |
| 5% error (15% Detection Limit) | | | | |
| 25 | 1.00e-03 | 3.17e-03 | 12.72 (-1.26) | 1.26 (+0.11) |
| 25 | 0.1 | 0.41 | 1.59 (-0.001) | 1.30 (-0.11) |
| 25 | 10 | 43.91 | 13.80 (+1.38) | 1.28 (-0.27) |
| 50 | 1.00e-03 | 1.35e-03 | 3.70 (-0.28) | 1.03 (-0.03) |
| 50 | 0.1 | 0.18 | 0.58 (-0.0003) | 1.14 (-0.27) |
| 50 | 10 | 10.17 | 2.82 (+0.23) | 1.23 (-0.17) |
| 100 | 1.00e-03 | 4.19e-04 | 0.62 (-0.03) | 1.06 (+0.06) |
| 100 | 0.1 | 0.07 | 0.27 (-0.0002) | 1.20 (-0.17) |
| 100 | 10 | 3.73 | 0.63 (+0.02) | 1.53 (-0.07) |
| 10% error (30% Detection Limit) | | | | |
| 50 | 1.00e-03 | 2.88e-03 | 12.68 (-1.23) | 1.34 (+0.11) |
| 50 | 0.1 | 0.31 | 1.53 (+0.002) | 1.30 (-0.04) |
| 50 | 10 | 39.29 | 13.67 (+1.40) | 1.29 (-0.24) |
| 100 | 1.00e-03 | 1.25e-03 | 3.44 (-0.26) | 1.00 (-0.03) |
| 100 | 0.1 | 0.16 | 0.57 (-0.001) | 0.92 (-0.32) |
| 100 | 10 | 9.06 | 3.79 (+0.27) | 0.97 (-0.16) |

Table 1. Precision and bias of potency metrics in Hill equation models. Equation (1) in the Methods was used to simulate 10,000 fifteen-point concentration-response profiles with the specified error for true values of $|RMAX|$ and AC_{50} , where the hill slope h was taken to be 1 and R_0 was specified as 0. \log_{10} Precision is shown with \log_{10} Bias in parentheses. Only data sets with true responses above the assay detection limit were evaluated.

| Data Set Parameters | | | | | Estimator | |
|---------------------------------|-------------------|-------------------|------------------|-----------|---------------|--------------|
| True $ RMAX $ | True $AC_{50(G)}$ | True $AC_{50(L)}$ | True POD_{WES} | True Peak | AC_{50} | POD_{WES} |
| 5% error (15% Detection Limit) | | | | | | |
| 25 | 0.001 | 0.1 | 3.23e-03 | 20.5 | 2.41 (+2.71) | 1.44 (+0.12) |
| 25 | 0.001 | 10 | 3.17e-03 | 24.5 | 7.48 (+2.03) | 1.32 (+0.11) |
| 25 | 0.1 | 10 | 0.36 | 20.5 | 3.28 (-0.82) | 1.37 (-0.07) |
| 50 | 0.001 | 0.1 | 1.43e-03 | 40.95 | 0.85 (+2.51) | 1.03 (-0.04) |
| 50 | 0.001 | 10 | 1.35e-03 | 49.0 | 6.42 (+1.95) | 1.03 (-0.03) |
| 50 | 0.1 | 10 | 0.18 | 40.95 | 0.88 (-0.51) | 1.14 (-0.27) |
| 100 | 0.001 | 0.001 | 5.80e-04 | 24.3 | 1.32 (+1.15) | 0.86 (+0.05) |
| 100 | 0.001 | 0.1 | 4.27e-04 | 81.9 | 0.37 (+2.48) | 1.07 (+0.05) |
| 100 | 0.001 | 10 | 4.19e-04 | 98.0 | 6.44 (+1.98) | 1.07 (+0.06) |
| 100 | 0.1 | 0.1 | 0.08 | 25.0 | 19.78 (-0.29) | 0.95 (-0.20) |
| 100 | 0.1 | 10 | 0.07 | 81.9 | 0.37 (-0.48) | 1.21 (-0.16) |
| 100 | 10 | 10 | 9.47 | 24.3 | 1.36 (-1.17) | 0.94 (-0.42) |
| 10% error (30% Detection Limit) | | | | | | |
| 50 | 0.001 | 0.1 | 2.55e-03 | 40.95 | 2.60 (+2.72) | 1.44 (+0.17) |
| 50 | 0.001 | 10 | 2.89e-03 | 49.0 | 7.33 (+1.96) | 1.32 (+0.12) |
| 50 | 0.1 | 10 | 0.31 | 40.95 | 3.47 (-0.80) | 1.36 (-0.04) |
| 100 | 0.001 | 0.1 | 1.37e-03 | 81.9 | 0.87 (+2.51) | 1.00 (-0.06) |
| 100 | 0.001 | 10 | 1.25e-03 | 98.0 | 6.42 (+1.95) | 1.00 (-0.03) |
| 100 | 0.1 | 10 | 0.16 | 81.9 | 0.87 (-0.51) | 0.94 (-0.32) |

Table 2. Precision and bias of potency metrics in gain-loss models. Equation (2) in the Methods was used to simulate data sets with the specified error. \log_{10} Precision is shown with \log_{10} Bias in parentheses. Only data sets with true responses above the assay detection limit were evaluated.

Figure 3 shows concentration-response profiles for four representative compounds that are tested once in each of the three experimental runs. Estradiol valerate is a synthetic ester of the positive control compound 17 β -estradiol and is consistently ranked within the top ten compounds based on *WES*. The corresponding potency value for this compound (POD_{WES}) is assigned a value of “less than the lowest tested concentration” in each run. Gestrinone is a synthetic hormone that elicits an agonistic response of $0.05 \pm 0.03 \mu\text{M}$ (across runs) in

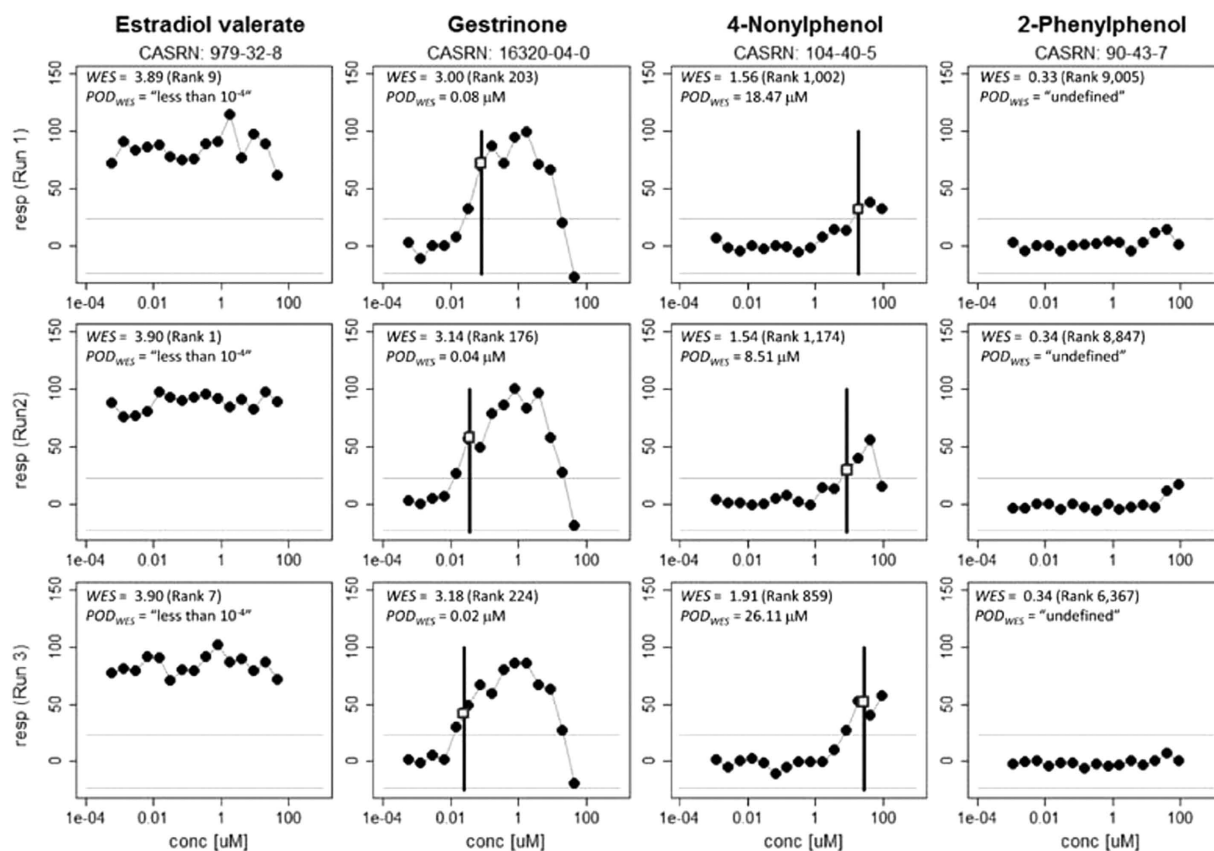


Figure 3. POD_{WES} values for example concentration-response profiles from the Tox21 Phase II BG1 ER alpha agonist data set. A total of four compounds are shown, with one compound in each column. Each row represents the concentration-response relationship from a separate experimental run. The ranking of each WES score out of all 11,776 compounds in the experimental run is given in parentheses. Gray lines indicate the 3σ assay detection limit.

this experiment, and is ranked in the top 250 compounds based on WES in each run shown here. As shown in Fig. 3, the response profile for this compound is better represented by the “gain-loss” model than the Hill model, perhaps due to cytotoxic effects at the greater concentrations. The next compound, 4-Nonylphenol, has previously been shown to act as an agonist of the estrogen receptor alpha in MCF7 breast cancer cells²⁴. This compound is ranked within the top 1,200 profiles based on WES and has a corresponding *in vitro* potency of $17.7 \pm 8.8 \mu\text{M}$. Finally, the biocide 2-Phenylphenol does not show detectable activity in the assay in any experimental run and is therefore ranked very low based on the WES score in each case. The potency of this compound is assigned the value of “undefined.”

The reproducibility of the potency estimates in this data set was evaluated by calculating \log_{10} potency differences between intra-assay duplicates and inter-assay duplicates interrogated across experimental runs. It is expected that duplicated compounds will have a \log_{10} mean ratio of 0, which corresponds to a mean ratio of 1.00 on the natural scale. All duplicated chemicals that had at least one observed response above the assay detection limit were included in the analysis. As shown in Fig. 4, there is less variation in \log_{10} potency differences for intra-array duplicates compared to inter-array duplicates as assessed by the median absolute deviation from zero. The dispersion of \log_{10} potency differences is noticeably greater for the AC_{50} value compared to POD_{WES} , indicating that AC_{50} values are less reproducible potency estimates in this experiment.

Discussion

High-throughput screening of compounds for biological activity can play a fundamental role in the advancement of drug discovery²⁵ and in efforts to transform toxicology from a mostly observational science into a predictive science²⁶. Large-scale qHTS data analyses typically proceed by fitting the Hill equation⁹ to the data and utilizing the AC_{50} value as an estimate for compound potency. However, the uncertainty (e.g., confidence intervals) of these nonlinear parameter estimates can be extremely large and potentially limit the reproducibility of results obtained from qHTS studies¹⁰. A new procedure is proposed here to estimate compound potency based on locating the maximum rate of change in weighted entropy. This approach (POD_{WES}) provides more precise estimates of potency than typically obtained by nonlinear parameter estimation from the Hill model.

Regardless of the level of error used to simulate the concentration-response curves, under most circumstances potency measures examined here were subject to empirical confidence interval widths spanning at least one order of magnitude. However, the CIW for AC_{50} estimates extended to greater than 13 orders of magnitude for

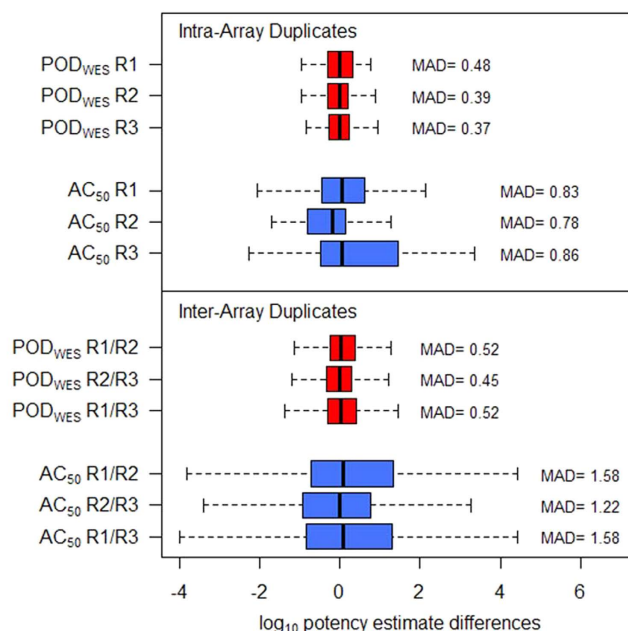


Figure 4. The distribution of \log_{10} potency differences (AC_{50} or POD_{WES}) for intra-array duplicate comparisons (within experimental run R1, R2 and R3) and inter-array duplicate comparisons (between two experimental runs R1/R2, R2/R3 and R1/R3). In the absence of experimental artifacts, it is expected that the \log_{10} potency differences would have a median value of zero, which corresponds to a potency ratio of one on the natural scale. The median absolute deviate from zero (MAD) is indicated on the figure for each set of \log_{10} potency differences.

low efficacy compounds at $|RMAX| = 25\%$ (see Table 1). Even so, the CIW of POD_{WES} was less than 1.53 orders of magnitude (less than 34-fold) for data simulated from a Hill equation model or a “gain-loss” model. The bias in POD_{WES} estimates was less than 2.7-fold (or $|\log_{10}Bias| \leq 0.42$), and usually less than 1.5-fold. AC_{50} estimates showed less bias than POD_{WES} for Hill model curves generated with two clearly defined asymptotes, but bias was much greater when the data was generated under other conditions.

Across-run comparisons of potency can be more variable than within-run comparisons (see Fig. 4). However, high-throughput screening responses can be affected by both random and systematic error, and run-to-run variability should be not ignored²⁷. Obtaining experimental replicates can increase the precision of the potency estimates²⁸ and the interpretation of POD_{WES} may be improved by focusing on robust assays with good agreement between compound measurements²⁹ and using appropriate signal curation⁶. If potencies are only desired from a pre-specified functional form (e.g., the Hill model), a two-step procedure can be used to (1) find response profiles that are active according to a robust analysis framework designed to detect the desired trend³⁰ and then (2) estimate potencies from the returned profiles.

The repeatability of AC_{50} estimates can be extremely small for compounds with low efficacy or for situations in which one of the asymptotes cannot be established¹⁰. Furthermore, the assay detection limit can impact the precision of potency estimation. Using 3σ of the negative controls as a detection limit is a common practice in qHTS studies^{1,6,28} and the 3σ value performed optimally in our simulation study across a range of σ values when considering bias, precision and the number of profiles with estimable potency values according to the Hill model (Fig. S2) and the “gain-loss” model (Fig. S3). Selective elimination of influential observations will not overcome these issues and may introduce bias because the true concentration-response relationship cannot be known in advance. Difficulties in characterizing the uncertainty of potency estimates derived from pre-specified models may be compounded when response profiles deviate from monotonicity or the incorrect model is employed for nonlinear curve fitting.

Each compound in a qHTS assay can be expected to have a distinctive set of parameters governing its response behavior. However, the approach proposed here to estimate potency using POD_{WES} does not rely on a pre-specified concentration-response pattern, can be applied to complex response patterns without respect to the direction of response and naturally accommodates missing data into the estimation framework.

Methods

This section describes the procedures used to estimate compound potency based on maximizing the rate of change in weighted entropy. Data sets are simulated based on the Hill equation in order to evaluate the precision and bias of estimated potencies across a range of parameter space characterizing qHTS studies. In addition, the new potency measure is applied to an experimental data set assaying for estrogen receptor agonist activity from phase II of Tox21.

Description of simulated data. Similar to previous studies^{7,10,30}, concentration-response data sets were simulated using the logistic form of the four-parameter Hill equation model,

$$R_i = R0 + \frac{(RMAX - R0)}{1 + 10^{\{-h[\log_{10}C_i - \log_{10}AC_{50}]\}}} + ERROR \quad (1)$$

where R_i is a normalized response representing a percentage of the positive control activity at concentration C_i . $RMAX$ is the maximal response, $R0$ is the minimal response, AC_{50} is the concentration of half-maximal response, h affects the shape of the curve and $ERROR$ is residual error of the model. The logarithm in equation (1) ensures that back-calculated estimates of AC_{50} obtained from $\log_{10}AC_{50}$ are restricted to positive values¹⁰. The concentrations (C_i) are based on equivalent \log_{10} concentration spacing ranging from 0.0001 to 100 μ M for fifteen-point concentration-response curves. The values of $RMAX$ and AC_{50} were set to (25, 50, 100% of positive control activity) and (10^{-3} , 10^{-1} , 10 μ M), respectively, for a total of 9 different data sets. The $R0$ parameter was set to 0 and h was set to 1. Other data sets were simulated using a “gain-loss” model defined as the product of two Hill equation models,

$$R_i = RMAX \left(\frac{1}{1 + 10^{\{\log_{10}AC_{50(G)} - \log_{10}C_i\}}} \right) \left(\frac{1}{1 + 10^{\{\log_{10}C_i - \log_{10}AC_{50(L)}\}}} \right) + ERROR \quad (2)$$

where $RMAX$ now represents a shared upper asymptote, both bottom asymptotes equal 0, $AC_{50(G)}$ is the concentration of half-maximal response in the gain direction and $AC_{50(L)}$ is the concentration of half-maximal response in the loss direction¹⁵. Similar to simulations performed using Eqn. (1), the values of $RMAX$, $AC_{50(G)}$ and $AC_{50(L)}$ were set to (25, 50, 100%), (10^{-3} , 10^{-1} , 10 μ M) and (10^{-3} , 10^{-1} , 10 μ M), respectively, for a total of 27 different data sets. However, only 12 of these data sets, for which the maximum response (or Peak Response) exceeded the specified detection limit, were included in the analyses here. Residual errors for equations (1) and (2) were modeled as $ERROR \sim N(0, \sigma^2)$ with $\sigma = 5\%$ or 10%, where σ is related to the percent of negative control responses producing variation levels often seen in Tox21 Phase II assays⁶. Unless otherwise noted, the assay detection limit is taken to be 3σ , a typical detection limit in HTS studies^{1,6,28}. A total of 10,000 simulated substances ($RMAX = 25\%$, 50%, or 100% of positive control activity) were simulated for each data set.

Description of estrogen receptor agonist data set. We acquired qHTS data involving approximately 10,000 compounds that were screened for estrogen receptor alpha agonist activity²³. This screen utilized an endogenous full length estrogen receptor alpha (BG1 cell line) with a luciferase reporter gene producing a single-channel readout²³. A total of 15 concentrations were evaluated with concentrations typically ranging from $\sim 10^{-3}$ μ M to ~ 78 μ M. As part of phase II of Tox21, the library is screened three times with compounds located in different well positions during each experimental run⁴. The raw plate reads were normalized using the positive and negative control wells and subsequently corrected for row, column, and plate effects using linear interpolation²³. Hill equation parameter estimates and activity calls were determined as described previously³⁰. In order to assess within-run reproducibility, a set of 88 broadly active duplicates were deliberately included in the Tox21 Phase II 1,536-well assay plates. Concentration-response patterns in this experimental data set encompass many different types of patterns which may deviate substantially from sigmoidal profiles.

Weighted entropy score. The weighted entropy score provides a measure of average relative activity across a concentration-response profile⁷. Briefly, the response vector for a given substance $\mathbf{R}_N = (R_1, \dots, R_N)$ contains an observed response R_i for each of N concentrations, where R_i corresponds to the response at the i th concentration, C_i . The relative response at C_i is defined as.

$$p_i = |R_i| / \sum_{i=1}^N |R_i| \quad (3)$$

where $p_i \geq 0$ and $\sum_{i=1}^N p_i = 1$. The relative responses p_i define a probability mass distribution based on the magnitude of R_i , where R_i may be positive or negative for activation or inhibition, respectively⁷. The entropy of R_i , or surprisal of the i th event, is defined as

$$H(R_i) = -\log_2 p_i \quad (4)$$

where the units of information are in bits. The weighted average entropy across the response profile takes into account the extent of each response compared to the detection limit of the assay. The weighted entropy score (WES) of a substance across N concentration levels is given by the expression

$$WES = -\sum_{i=1}^N w_i p_i \log_2 p_i \quad (5)$$

where $WES \geq 0$ and by convention $0 \log_2 0 = 0$ since $\lim_{p \rightarrow 0} p \log_2 p = 0$. When every response value is greater than or equal to the assay detection limit, all $w_i = 1$ so that WES is the same as Shannon entropy (i.e., $WES = H = -\sum p_i \log_2 p_i$). However, when R_i values are less than the assay detection limit of 3σ the weights w_i are defined as the ratio of the surprisal frequency for a relative response within the assay noise region (i.e., $-p_{i,noise} \log_2 p_{i,noise}$ where $p_{i,noise} = p_i/3\sigma$) divided by the uncorrected surprisal frequency (i.e., $-p_i \log_2 p_i$), or $w_i = -p_{i,noise} \log_2 p_{i,noise} / -p_i \log_2 p_i = (p_i/3\sigma) \log_2 (p_i/3\sigma) / (p_i \log_2 p_i)$. Larger values of WES indicate more detectable responses

across concentration levels⁷. The entropy at the *k*th tested concentration (H_k or WES_k) is computed by considering only the responses $\mathbf{R}_k = (R_1, \dots, R_k)$ that are observed within the full concentration-response profile \mathbf{R}_N .

The POD Approach to Estimate Potency. We define the profile-specific potency (denoted Point of Departure, POD) as the concentration along the response profile at which the magnitude of the rate of change in WES is greatest. This maximum rate of change defines the potency regardless of the direction of change, i.e., irrespective of whether the chemical is an activator or inhibitor. The rate of change in WES across the profile is computed as the derivative of WES with respect to concentration, or $\frac{dWES}{dC}$, where concentration is based on \log_{10} units. In mathematical terms, POD_{WES} is located at the concentration with the maximal value of $\left| \frac{dWES}{dC} \right|$ where $\frac{d^2WES}{dC^2}$ is equal to zero and either (a) $\frac{d^2WES}{dC^2}$ changes sign from positive to negative (for activation) or (b) $\frac{d^2WES}{dC^2}$ changes sign from negative to positive (for inhibition) according to “The First Derivative Test”³¹. We compute the derivatives of WES using finite difference calculus, a mathematical procedure based on a Taylor series procedure that provides difference formulas for a grid sampled at discrete data points³². If there are no detectable responses in the profile, the potency is declared “undefined”. However, if potency cannot be estimated within the observed response profile but a detectable response is found within data region, finite difference calculus is used to predict the assay response beyond the tested concentration range. This extrapolation continues until POD_{WES} is quantitatively estimated or designated “less than C_1 ” for profiles that have substantial activity at the lowest observed concentration but no quantitative estimate is obtainable (see Fig. 1). No data points located within the detection window are extrapolated outside of the noise region. The estimation of POD_{WES} is described in greater detail in the Supplementary Information.

Evaluating potency estimates. AC_{50} estimates were determined according to Shockley³⁰. The POD_{WES} approach was described above and presented in Fig. 1. The precision of each potency estimator was investigated by calculating the empirical 95% confidence interval widths (2.5th percentile – 97.5th percentile) of the \log_{10} transformed estimates within a generated data set. Bias was calculated by subtracting the “true” value θ of potency estimator U (e.g., AC_{50} or POD_{WES} obtained from profiles simulated with ERROR = “0%”) from the mean of the estimated values of U according to $\frac{1}{n} \sum_{k=1}^n U_k - \theta$. Evaluations of potency estimates are performed using the \log_{10} transformation so that the distributions of potencies better approximate a normal distribution with constant error³³. All computations were performed in the statistical software R³⁴.

References

- Inglese, J. *et al.* Quantitative high-throughput screening: a titration-based approach that efficiently identifies biological activities in large chemical libraries. *Proc Natl Acad Sci USA* **103**, 11473–11478 (2006).
- Reinhold, W. C. *et al.* Using drug response data to identify molecular effectors, and molecular “omic” data to identify candidate drugs in cancer. *Hum Genet* **134**, 3–11 (2015).
- Zhu, H. *et al.* Big data in chemical toxicity research: the use of high-throughput screening assays to identify potential toxicants. *Chem Res Toxicol* **27**, 1643–1651 (2014).
- Tice, R. R., Austin, C. P., Kavlock, R. J. & Bucher, J. R. Improving the human hazard characterization of chemicals: a Tox21 update. *Environ Health Perspect* **121**, 756–765 (2013).
- Beam, A. & Motsinger-Reif, A. Beyond IC : Towards Robust Statistical Methods for Association Studies. *J Pharmacogenomics Pharmacoproteomics* **5**, 1000121 (2014).
- Hsieh, J. H., Sedykh, A., Huang, R., Xia, M. & Tice, R. R. A Data Analysis Pipeline Accounting for Artifacts in Tox21 Quantitative High-Throughput Screening Assays. *J Biomol Screen* **20**, 887–897 (2015).
- Shockley, K. R. Using weighted entropy to rank chemicals in quantitative high-throughput screening experiments. *J Biomol Screen* **19**, 344–353 (2014).
- Thomas, R. S. *et al.* A comprehensive statistical analysis of predicting *in vivo* hazard using high-throughput *in vitro* screening. *Toxicol Sci* **128**, 398–417 (2012).
- Hill, A. V. The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *J Physiol* **40**, 4–7 (1910).
- Shockley, K. R. Quantitative high-throughput screening data analysis: challenges and recent advances. *Drug Discov Today* **20**, 296–300 (2015).
- Bergeron, C., Moore, G., Krein, M., Breneman, C. M. & Bennett, K. P. Exploiting domain knowledge for improved quantitative high-throughput screening curve fitting. *J Chem Inf Model* **51**, 2808–2820 (2011).
- Fujii, Y., Narita, T., Tice, R. R., Takeda, S. & Yamada, R. Isotonic Regression Based-Method in Quantitative High-Throughput Screenings for Genotoxicity. *Dose Response* **13**, doi: 10.2203/dose-response.13-045.Fujii (2015).
- Conolly, R. B. & Lutz, W. K. Nonmonotonic dose-response relationships: mechanistic basis, kinetic modeling, and implications for risk assessment. *Toxicol Sci* **77**, 151–157 (2004).
- Peddada, S. D. & Haseman, J. K. Analysis of nonlinear regression models: a cautionary note. *Dose Response* **3**, 342–352 (2005).
- EPA (Environmental Protection Agency). ToxCast™ Data. The ToxCast Analysis Pipeline: An R package for processing and modeling chemical screening data. https://www.epa.gov/sites/production/files/2015-08/documents/pipeline_overview.pdf (2016) (Date of access: March 25, 2016).
- Crump, K. S. A new method for determining allowable daily intakes. *Fundam Appl Toxicol* **4**, 854–871 (1984).
- Woutersen, R. A., Jonker, D., Stevenson, H., te Biesebeek, J. D. & Slob, W. The benchmark approach applied to a 28-day toxicity study with Rhodorsil Silane in rats. the impact of increasing the number of dose groups. *Food Chem Toxicol* **39**, 697–707 (2001).
- Shannon, C. E. A mathematical theory of communication. *Bell Syst Techn J.* **27**, 1–55 (1948).
- Fuhrman, S. *et al.* The application of shannon entropy in the identification of putative drug targets. *Biosystems* **55**, 5–14 (2000).
- Schug, J. *et al.* Promoter features related to tissue specificity as measured by Shannon entropy. *Genome Biol* **6**, R33 (2005).
- Zhang, Y. *et al.* QDMR: a quantitative method for identification of differentially methylated regions by entropy. *Nucleic Acids Res* **39**, e58 (2011).
- Cover, T. M. & Thomas, J. A. *Elements of information theory*. (John Wiley & Sons, 1991).
- Huang, R. *et al.* Profiling of the Tox21 10K compound library for agonists and antagonists of the estrogen receptor alpha signaling pathway. *Sci Rep* **4**, 5664 (2014).
- Vivacqua, A. *et al.* The food contaminants bisphenol A and 4-nonylphenol act as agonists for estrogen receptor alpha in MCF7 breast cancer cells. *Endocrine* **22**, 275–284 (2003).

25. Macarron, R. *et al.* Impact of high-throughput screening in biomedical research. *Nat Rev Drug Discov* **10**, 188–195 (2011).
26. Collins, F. S., Gray, G. M. & Bucher, J. R. Toxicology. Transforming environmental health protection. *Science* **319**, 906–907 (2008).
27. Kevorkov, D. & Makarenkov, V. Statistical analysis of systematic errors in high-throughput screening. *J Biomol Screen* **10**, 557–567 (2005).
28. Malo, N., Hanley, J. A., Cerquozzi, S., Pelletier, J. & Nadon, R. Statistical practice in high-throughput screening data analysis. *Nat Biotechnol* **24**, 167–175 (2006).
29. Ilouga, P. E. & Hesterkamp, T. On the prediction of statistical parameters in high-throughput screening using resampling techniques. *J Biomol Screen* **17**, 705–712 (2012).
30. Shockley, K. R. A three-stage algorithm to make toxicologically relevant activity calls from quantitative high throughput screening data. *Environ Health Perspect* **120**, 1107–1115 (2012).
31. Marsden, J. & Weinstein, A. *Calculus I*. (Springer-Verlag New York Inc., 1985).
32. Lynch, D. R. *Numerical partial differential equations for environmental scientists and engineers*. (Springer, 2005).
33. Altman, D. G. & Bland, J. M. Statistics notes: the normal distribution. *BMJ* **310**, 298 (1995).
34. R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria, URL <http://www.R-project.org/>, 2012).

Acknowledgements

I thank Dr. Raymond Tice (Biomolecular Screening Branch, NIEHS) and Dr. Grace Kissling (Biostatistics and Computational Biology Branch, NIEHS) for reviewing the manuscript and providing helpful suggestions. This work was supported [in part] by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (ZIA ES102865).

Author Contributions

K.R.S. designed the study, analyzed the data, wrote the manuscript and edited the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The author declares no competing financial interests.

How to cite this article: Shockley, K. R. Estimating Potency in High-Throughput Screening Experiments by Maximizing the Rate of Change in Weighted Shannon Entropy. *Sci. Rep.* **6**, 27897; doi: 10.1038/srep27897 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>