The Estimation of Absolute IC_{50} and Its 95% Confidence Interval

Qin Liu
Molecular and Cellular Oncogenesis Program, The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104.

Abstract

Inhibitory concentration 50% (IC_{50}) is commonly used to compare the potency of drugs in pre-clinical anticancer drug screening study. IC_{50} represents the concentration of a drug that is required to inhibit 50% of cancer cells in the *in vitro* experiment. The IC_{50} of a drug is usually estimated using four-parameter logistic (4PL) regression analysis. When comparing IC_{50}s between drugs, it needs to clear if the relative IC_{50} or the absolute IC_{50} was estimated. In practice, different drug may reach different maximum and minimum inhibition rates along the tested dose levels. By taking into account such variation in model fitting, the 4PL model estimated IC_{50} is the relative IC_{50}, which is not comparable among tested drugs. This paper has proposed an approach for the calculation of the comparable IC_{50} (absolute IC_{50}) and provided a formula for variance estimation with the delta method. After obtaining the absolute IC_{50} and its variance, the 95% confidence interval of the absolute IC_{50} is estimated and readily suited to comparisons among drugs.

Key Words: dose-response, IC_{50}, four-parameter logistic regression model, delta method

1. Introduction

In pre-clinical anticancer drug screening research, *in vitro* dose-response experiments are usually used first to identify drugs with promising high potency in inhibiting cancer cells. With dose-response data from a particular drug, inhibitory concentration 50% (IC_{50}) can be estimated and applied to assess the magnitude of potency for the studied drug. IC_{50} indicates the concentration (dose) of the studied drug that is needed to inhibit cancer cells by 50%.

To estimate IC_{50} of a particular drug, the four-parameter logistic (4PL) regression model, expressed as equation (1), is commonly used [1-5].

$$y = \frac{U-L}{1+10^{(\log[IC50]-\log[D_y])m}} + L \qquad (1)$$

In equation (1), the four parameters of 4PL regression model include IC_{50}, the Hill slope (m, refers to the steepness of a dose-response curve), the maximum (U) and minimum (L) response (the percent inhibition reflected by the top and bottom asymptotes of a dose-response curve shown as Figure 1). D_y is the dose of the studied drug that can induce a response (inhibition rate) as big as y. 4PL regression model is an alternative format of equation (2), the well-known Hill equation [6].

$$y = \frac{U-L}{1+\left(\frac{IC50}{D_y}\right)^m} + L \quad (2)$$

In a dose-response experiment, a series dose of a studied drug from low to high are to be tested. The responses corresponding to tested dose levels are usually normalized as the percentage of cancer cells that were killed or inhibited by the given drug. Theoretically, the normalized minimum response is 0%, and the maximum response is 100% for the controls. Ideally, for the tested drugs, the lower levels of the tested doses generate very low responses that close to 0%, and the higher levels of the tested doses tend to reach the 100% response. A dose-response curve as such case is demonstrated in Figure 1.

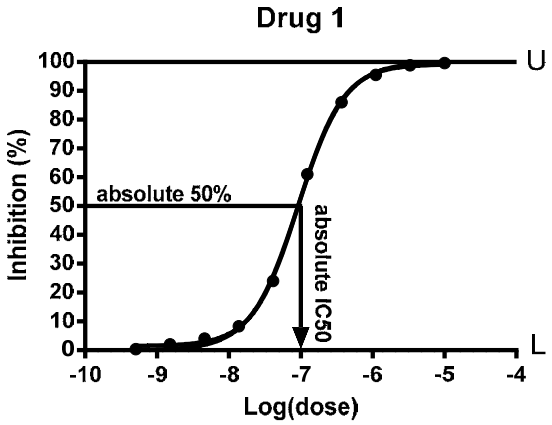

Figure 1: Dose-response curve of drug 1

Over the tested dose levels, if the studied drug's dose-response curve has the minimum response closing to 0% and the maximum response reaching 100%, its IC_{50} is called as absolute IC_{50} [7,8] since it is the dose inhibiting absolute 50% of cancer cells (Figure 1). One drug's absolute IC_{50} can be compared with another drug's absolute IC_{50} directly. A drug with a lower absolute IC_{50} indicates higher potency than another drug with a greater IC_{50}.

It is not uncommon, however, to see different minimum and maximum responses among the dose-response curves from studied drugs in a drug screening study. Some drugs may never be able to reach 100% response no matter how high dose levels were given [7,8]. The dose-response data shown in Figure 2 is such a case. Usually, most drugs almost generate no response at very low dose levels, while some drugs may reach certain response level far beyond 0% at similar low dose levels [2,4]. By taking into account these variations in model fitting, the IC_{50} calculated using the 4PL regression model is actually the relative IC_{50} [7,8], which is the dose that generates the half level of response between the minimum and maximum responses where the studied drug could reach in the experiment. The relative IC_{50} should not be used directly for comparisons across drugs. In order to compare the potencies across the tested drugs in the drug screening study, the absolute IC_{50} is comparable and should be considered for the application. Example data shown in Figures 1 and 2 illustrate the difference between the absolute IC_{50} and the relative IC_{50}. In Figure 1, the drug 1's dose-response curve shows that its U=100%, L=0%, and the absolute IC_{50} is around -7 ($\log_{10}$ scale of dose). In Figure 2, the drug 2's U is about 69%, L stays as 0%, and the absolute IC_{50} was around -6.5 ($\log_{10}$ scale of dose). The 4PL regression analysis would report a IC_{50} about -7 in $\log_{10}$ scale of dose, which is

the relative IC_{50}, indicating the dose that generates the half of the level response (around 35% inhibition for drug 2) between the drug's U and L.

A simply comparison based on a relative IC_{50} would lead wrong conclusion. For example, drug 2's relative IC_{50} equals to drug 1's absolute IC_{50}. It will be wrong to conclude that drug 1 has same potency as drug 2. To appropriately compare drugs 1 and 2, their absolute IC_{50} should be compared and the conclusion is that drug 1 shows stronger potency than drug 2 since drug 2's absolute IC_{50} is higher than drug 1's absolute IC_{50}.

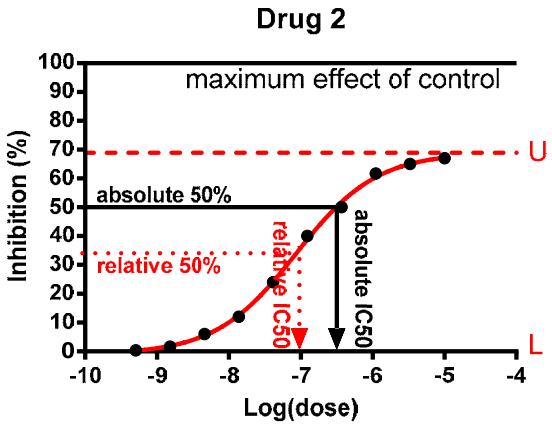

Figure 2: Dose-response curve of drug 2

In drug-screening studies, the absolute IC_{50} and corresponding 95% confidence interval (CI) for each drug need to be estimated for appropriate comparison. Nevozhay [8] developed Cheburator software to calculate the absolute IC_{50} and 95% CI using bootstrapping method for experiments from 96-well plates with triplicate at 4 concentrations for each tested drug. Without established software for a set of specifically designed experiments, it is not easy to perform bootstrap by regular lab without professional statistical support. Another problem for bootstrap is that the procedure of calculation is time consuming when the number of tested drugs are large.

To solve above difficulties in practice, an approach is proposed in Methods section of this article for estimating the absolute IC_{50} with 95% CI. Its application is shown in the section of Examples and Results.

2. Methods

The statistical function such as non-linear regression for dose-response in GraphPad Prism (GraphPad Software, Inc.), nl log4 in STATA (StataCorp LP, 4905 Lakeway Drive, College Station, TX 77845, USA) and the drfit or drc packages in the R Statistical Environment are popular tools to calculate the relative IC_{50} with 4PL regression analysis. With the estimated four parameters of the 4PL regression model, the absolute IC_{50} and its 95% CI can be estimated using following proposed approach.

Based on equation (2), the dose D_y of the tested drug that generates the absolute response as big as y can be expressed as equation (3).

$$D_y = IC50 * (\frac{y-L}{U-y})^{\frac{1}{m}} \quad (3)$$

$IC50$ in equation (3) is the relative IC_{50}. The absolute IC_{50} is represented as D_{50} in equation (3) when y=50%. The D_{50} can be estimated as

$$D_{50} = IC50 * \left(\frac{50-L}{U-50}\right)^{\frac{1}{m}} = \widehat{IC50} * \left(\frac{50-\hat{L}}{\hat{U}-50}\right)^{\frac{1}{\hat{m}}} \quad (4)$$

Where $\widehat{IC50}$, $\hat{L}$, $\hat{U}$, and $\hat{m}$ are the estimates obtained from the 4PL model. With base 10 log-transformation, equation (4) is turned into equation (5).

$$\log(D_{50}) = \frac{\log\left(\frac{50-L}{U-50}\right)}{m} + \log(IC50) \ = \frac{\log\left(\frac{50-\hat{L}}{\hat{U}-50}\right)}{\hat{m}} + \log\left(\widehat{IC50}\right) \quad (5)$$

Based on equation (5), with the delta method [11], a formula for estimating the variance of $\log(D_{50})$, expressed as $Var(\log(D_{50}))$, is deducted as below:

$$\begin{aligned} Var(\log(D_{50})) &= Var\left(\frac{\log\left(\frac{50-\hat{L}}{\hat{U}-50}\right)}{\hat{m}} + \log\left(\widehat{IC50}\right)\right) \\ &= Var\left(\frac{\log(50-\hat{L})}{\hat{m}} - \frac{\log(\hat{U}-50)}{\hat{m}} + \log\left(\widehat{IC50}\right)\right) \\ &\approx Var\left(\frac{\log(50-\hat{L})}{\hat{m}}\right) + Var\left(\frac{\log(\hat{U}-50)}{\hat{m}}\right) + Var\left(\log\left(\widehat{IC50}\right)\right) \\ &\approx \frac{1}{\hat{m}^2} * \frac{1}{((50-\hat{L})*ln10)^2} * Var(L) + \frac{(\log(50-\hat{L}))^2}{\hat{m}^4} * Var(m) \\ &+ \frac{1}{\hat{m}^2} * \frac{1}{((\hat{U}-50)*ln10)^2} * Var(U) + \frac{(\log(\hat{U}-50))^2}{\hat{m}^4} * Var(m) \\ &+ Var(\log(IC50)) \end{aligned} \quad (6)$$

where $ln10$ is the natural logarithm of 10. $\mathrm{Var}(\log(IC50))$, $Var(m)$, $Var(L)$ and $Var(U)$ are the variances of the four parameters, which are the squared terms of the corresponding standard errors obtained from the 4PL regression analysis. If the equation (4) is natural log-transformed, the variance of $\ln(D_{50})$ can be estimated with equation (6) without the term of "$* ln10$".

Equation (6) provides an approximating estimate for the variance of the absolute IC_{50}. With this variance, the 95% CI of the absolute IC_{50} is then estimated using equation (7).

$$\log(D_{50}) - 1.96\sqrt{Var(\log(D_{50}))},\ \log(D_{50}) + 1.96\sqrt{Var(\log(D_{50}))} \quad (7)$$

Similar way can be used to estimate the variance of D_y corresponding to any y-level of absolute response. For example, the natural log-transformed dose that produces 25% response, $\ln(D_{25})$, is estimated as

$$\ln(D_{25}) = \frac{\ln\left(\frac{25-\hat{L}}{\hat{U}-25}\right)}{\hat{m}} + \ln\left(\widehat{IC50}\right)$$

and its variance, $Var(\ln(D_{25}))$ is estimated as below

$$Var(\ln(D_{25})) \approx \frac{1}{\hat{m}^2} * \frac{1}{((25-\hat{L}))^2} * Var(L) + \frac{(\ln(25-\hat{L}))^2}{\hat{m}^4} * Var(m) + \frac{1}{\hat{m}^2} * \frac{1}{((\hat{U}-25))^2} * Var(U) + \frac{(\ln(\hat{U}-25))^2}{\hat{m}^4} * Var(m) + Var(\ln(IC50))$$

3. Examples and Results

Table 1 are the data used to generate the dose-response curves in Figures 1 and 2. These data were simulated as an experiment with triplicates for each drug. They are used as examples to illustrate the above proposed approach for estimating the absolute IC_{50} and its 95% CI.

Table 1: Example data

dose (mol/L)	$\log_{10}$ (dose)	Drug 1 Inhibition (%)			Drug 2 Inhibition (%)		
0.00001	-5	100	100	99	66	68	67
3.31E-06	-5.48	99	98.9	98.6	65	64	66
1.12E-06	-5.95	95.5	96	95	60	63	62
3.72E-07	-6.43	87	85	86	50	50	50
1.23E-07	-6.91	60	61	62	40	39	41
4.07E-08	-7.39	25	23	24	25	23	24
1.38E-08	-7.86	10	8	7	13	11	12
4.57E-09	-8.34	4	5	3	6	5	7
1.51E-09	-8.82	2	1	3	1	2	2
5.13E-10	-9.29	0	0	1	0	1	0

From the 4PL regression models, the estimated four parameters are reported in Tables 2 and 3 for drugs 1 and 2, respectively. For each drug, the absolute IC_{50} is first estimated using a 4PL regression model (model 1) by restricting the top asymptote of the dose-response curve as U=100%, and the bottom asymptote as L=0%. Such restriction is acceptable for data like drug 1 whose maximum response closes to 100% and the minimum response closes to 0%. While for data like drug 2, such restrictions may decrease goodness of model fitting. The results from a 4PL regression model without such restrictions (model 2) were also reported for comparing models in goodness of fit. Model 1 provides the absolute IC_{50} directly. Model 2, however, obtains the relative IC_{50}.

For drug 1, the estimates of the four parameters from models 1 and 2 are almost the same. Bias-corrected Akaike Information Criterion (AICc), the statistics for goodness of fit, also indicates that both model 1 and model 2 have similar goodness of fit (Table 2). For data like drug 1, it will be easier to get appropriate estimate of the absolute IC_{50} and its 95% CI by using model 1. Per the output from model 1, drug 1's absolute IC_{50} (95% CI) is -7.038 (-7.054, -7.021) in $\log_{10}$ scale or 9.17E-08 (8.83E-08, 9.52E-08) mol/L.

Table 2: The estimates of four parameters (standard errors) from 4PL models for drug 1

Model	L (%)	U (%)	$\log_{10} IC50$	m	AICc
1	0 (-)	100 (-)	-7.038 (0.008)	1.316 (0.028)	26.3
2	1.463 (0.421)	99.21 (0.460)	-7.031 (0.008)	1.393 (0.035)	20.0

For drug 2, comparing the two models, the estimates of the four parameters are not even close. AICc indicates that Model 2 improves the goodness of fit a lot than model 1 (Table 3). Thus, the absolute IC_{50} estimated from model 1 is not precise. To get a precise estimate of the absolute IC_{50}, model 2 is used.

Table 3: The estimates of four parameters (standard errors) from 4PL models for drug 2

Model	L (%)	U (%)	$\log_{10} IC50$	m	AICc
1	0 (-)	100 (-)	-6.220 (0.060)	0.430 (0.027)	108.8
2	-1.030 (0.691)	68.95 (0.766)	-7.063 (0.022)	0.775 (0.032)	20.5

For drug 2, using results from model 2 and equations (5) to (7), the estimate of the absolute IC_{50} in $\log_{10}$ scale is -6.508, its corresponding variance is 0.0139, and estimated 95% CI is between -6.739 and -6.277. By anti-logarithm, the estimated absolute IC_{50} (95% CI) is 3.1E-07 (1.83E-07, 5.29E-07) mol/L. Comparing with drug 1's absolute IC_{50} and 95% CI, drug 2's absolute IC_{50} is significant higher.

4. Discussion

In the early stage of drug screening research, IC_{50} is one necessary statistics for the assessment of drug potency. Some software such as GraphPad Prism (GraphPad Software, Inc.), STATA (StataCorp LP, 4905 Lakeway Drive, College Station, TX 77845, USA) and R have functions for the 4PL regression analysis. The IC_{50} estimated by the 4PL regression analysis is actually the relative IC_{50} when the model estimated top or bottom plateaus of a dose-response curve are different from 100% or 0%. When the top and/or bottom plateaus are very different across the tested drugs, the relative IC_{50} estimated from the 4PL model is not suitable for the purpose of comparison. But with the four parameters estimated from the 4PL model, a comparable IC_{50}, the absolute IC_{50} with 95% CI can be further estimated using equations (5) to (7) proposed in this article.

Equations (5) to (7) are very easy to use in practice. For dose-response data that the minimum response is not close to 0% or the maximum response is not close to 100%, after fitting an appropriate 4PL model, equations (5) to (7) is readily suited to estimating the absolute IC_{50} and its 95% CI. Similar method can be applied to estimate inhibitory concentrations for other levels of absolute response and corresponding 95% CIs. This method is useful for drug screening studies.

References

1. Zhang L, He M, Zhang Y, Nilubol N, Shen M, and Kebebew E. Quantitative high-throughput drug screening identifies novel Classes of drugs with anticancer activity in thyroid cancer cells. J. Clin. Endocrinol. Metab 2012;97(3):E319–E328.
2. Mahida JP, Antczak C, DeCarlo D, Champ KG, Francis JH, Marr B, et al. A Synergetic screening Approach with companion effector for combination therapy: application to retinoblastoma. PLoS One 2013;8(3):e59156. doi:10.1371/journal.pone.0059156.
3. Ma W-Y, Hsiung L-C, Wang C-H, Chiang C-L, Lin C-H, Huang C-S, et al. A novel 96well-formatted micro-gap plate enabling drug response profiling on

primary tumour samples. Scientific Reports 2015;5:9656. doi:10.1038/srep09656.
4. Bharadwaj U, Eckols TK, Kolosov M, Kasembeli MM, Adam A, Torres D, et al. Drug-repositioning screening identified piperlongumine as a direct stat3 inhibitor with potent activity against breast cancer. Oncogene 2015;34(11):1341-1353. doi:10.1038/onc.2014.72.
5. Pozdeyev N, Berlinberg A, Zhou Q, Wuensch K, Shibata H, Wood,WM, et al. Targeting the NF-κB pathway as a combination therapy for advanced thyroid cancer. PLoS One 2015;*10*(8), e0134901. http://doi.org/10.1371/journal.pone.0134901.
6. Hill, A. V. The possible effects of the aggregation of themolecules of haemoglobin on its dissociation curves. J. Physiol 1910;40, iv–vii.
7. Sebaugh JL. Guidelines for accurate EC50/IC50 estimation. Pharm Stat. 2011;10(2):128-34. doi: 10.1002/pst.426.
8. Nevozhay D. Cheburator Software for Automatically Calculating Drug Inhibitory Concentrations from *In Vitro* Screening Assays. PLoS One 2014;9(9):e106186. doi:10.1371/journal.pone.0106186.
9. Ranke J. Fitting dose-response curves from bioassays and toxicity testing. R News 2006;3: 7–12.
10. Ritz C, Strebig J, Ritz MC. Package 'drc'. 2015. Available: https://cran.**r**-project.org/web/packages/drc/drc.pdf
11. Casella G, Berger RL. Statistical Inference. Thomson Learning, 2002. pp. 240-245.