A New Method with Flexible and Balanced Control of False Negatives and False Positives for Hit Selection in RNA Interference High-Throughput Screening Assays

XIAOHUA DOUGLAS ZHANG

The z-score method and its variants for testing mean difference are commonly used for hit selection in high-throughput screening (HTS) assays. Strictly standardized mean difference (SSMD) offers a way to measure and classify the short interfering RNA (siRNA) effects. In this article, based on SSMD, the authors propose a new testing method for hit selection in RNA interference (RNAi) HTS assays. This SSMD-based method allows the differentiation between siRNAs with large and small effects on the assay output and maintains flexible and balanced control of both the false-negative rate, in which the siRNAs with strong effects are not selected as hits, and the restricted false-positive rate, in which the siRNAs with weak or no effects are selected as hits. This method directly addresses the size of siRNA effects represented by the strength of difference between an siRNA and a negative reference, whereas the classic z-score method and t-test of testing no mean difference address whether the mean of an siRNA is exactly the same as the mean of a negative reference. This method can readily control the false-negative rate, whereas it is nontrivial for the classic z-score method and t-test to control the false-negative rate. Therefore, theoretically, the SSMD-based method offers better control of the sizes of siRNA effects and the associated false-positive and false-negative rates than the commonly used z-score method and t-test for hit selection in HTS assays. The SSMD-based method should generally be applicable to any assay in which the end point is a difference in signal compared to a reference sample, including those for RNAi, receptor, enzyme, and cellular function. (*Journal of Biomolecular Screening* 2007:000-000)

Key words: strictly standardized mean difference, d^+-probability, restricted false-positive rate, completely balanced error level, high-throughput screening, hit selection

INTRODUCTION

RNA INTERFERENCE (RNAi) OFFERS a safe and effective way of turning off a gene.[1,2] RNAi plays an important role in drug discovery. It may also be developed directly into drugs. In fact, RNAi has been seen as the 3rd class of drugs, after small molecules and proteins.[3,4] RNAi high-throughput screening (HTS) enables massive parallel gene silencing to identify genes associated with specific biological phenotypes. It is increasingly being used to reveal novel connections between genes and disease-relevant phenotypes.[5–10] Statistical methods for small-molecule HTS data have been described.[11–18] For quality control, the Z factor is broadly used in small-molecule HTS assays.[5,11] A more meaningful statistical parameter, strictly standardized mean difference (SSMD), has recently been proposed and evaluated for quality control in RNAi HTS assays.[19–21]

The most commonly used method for hit selection in HTS experiments is statistical significance (or p-value) from either the z-score method or t-test for testing no mean difference.[15–18,22–24] However, there are many issues with the classic statistical significance; thus, it has been criticized.[25–27] First, the p-value of testing no mean difference addresses whether the mean of a short interfering RNA (siRNA) is exactly the same as the mean of a negative reference. It does not measure the strength of difference between an siRNA and a negative reference directly. Second, the p-value of testing no mean difference is a function of both the sample size in the study and the magnitude of difference in the populations. In other words, the sample size and the magnitude of difference are indistinguishable in this statistical significance.[25] Third, the classic statistical significance of testing no mean difference does not control the false-negative and false-positive rates as well as scientists would like.[25,26]

Biometrics Research, Merck Research Laboratories, West Point, Pennsylvania.

Received Dec 16, 2006, and in revised form Jan 19, 2007. Accepted for publication Feb 1, 2007.

Journal of Biomolecular Screening XX(X); XXXX
DOI: 10.1177/1087057107300645

Both the z-score method and t-test of testing mean difference control the false-positive rate (namely, type I error), in which we conclude that $\mu_1 \neq \mu_2$, whereas, actually, $\mu_1 = \mu_2$. Another type of error is type II error, in which we conclude that $\mu_1 = \mu_2$, whereas, actually, $\mu_1 \neq \mu_2$. However, in practice, what we are really interested in is not whether an siRNA (or compound) has average inhibition/activation effects that are exactly the same as the negative reference. Instead, we are interested in the siRNAs or compounds with the largest magnitude of difference between the siRNAs and the negative reference. In the process of hit selection, we do not want to miss the siRNAs or compounds with true large effects.

To address these issues in hit selection caused by the use of classic statistical significance, SSMD has been proposed to measure the magnitude of difference between 2 compared groups.[19] SSMD can assess the siRNA effect represented by the magnitude of difference between an siRNA and a negative reference group. In addition, the link between SSMD and d^+-probability (the probability that the difference is positive) offers a clear interpretation of siRNA effects from a probability perspective.[19] In this article, based on SSMD, we further propose a new method for hit selection. In this method, we can first specify a targeted large magnitude of difference c and a small magnitude of difference c_2 and then maintain a flexible and balanced control of both the false-negative rate, in which the siRNAs with strong effects (namely, SSMD $\geq$ c) are not selected as hits, and the restricted false-positive rate, in which the siRNAs with weak or no effects (namely, SSMD $\leq c_2$) are selected as hits.

This article is organized as follows. First, we present briefly a pair of simple statistical parameters, SSMD and coefficient of variability in difference (CVD); their probability meaning; and statistical estimation and inference. Second, we elaborate on the new SSMD-based method for selecting hits with large effects. Finally, we discuss how to use this new method for hit selection in RNAi HTS experiments. In the follow-up article,[28] we will apply this pair of statistical parameters and this new SSMD-based method for hit selection in real primary RNAi HTS experiments.

A PAIR OF SIMPLE STATISTICAL PARAMETERS

SSMD and CVD

Suppose, in 2 populations with random values, the 1st population has mean μ_1 and variance σ_1^2 and the 2nd population has mean μ_2 and variance σ_2^2. The covariance between these 2 populations is σ_{12}. Let random variables P_1 and P_2 denote the 2 populations, respectively. Let D denote the difference between these 2 random variables, namely, $D = P_1 - P_2$. The standardized difference Z_D has mean μ_D and standard deviation σ_D. Let β and ω denote the SSMD and CVD, respectively. Then SSMD and CVD are defined as $\beta = \frac{\mu_D}{\sigma_D} = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}}$ and $\omega = \frac{\sigma_D}{\mu_D} = \frac{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}}{\mu_1 - \mu_2}$, respectively. If the 2 populations are independent, then $\beta = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}$ and $\omega = \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\mu_1 - \mu_2}$. If the 2 independent populations have equal variances (namely, $\sigma_1^2 = \sigma_2^2 = \sigma^2$), then $\beta = \frac{\mu_1 - \mu_2}{\sqrt{2\sigma^2}}$ and $\omega = \frac{\sqrt{2\sigma^2}}{\mu_1 - \mu_2}$. SSMD is the ratio of mean to standard deviation of the random variable representing the difference between the 2 populations, whereas CVD is the coefficient of variation of this random variable. Clearly, CVD is the reciprocal of SSMD. However, CVD $= \infty$ when $\mu_1 = \mu_2$, which may make CVD less favorable than SSMD in some situations. The value of SSMD denotes the size of the ratio of mean to standard deviation of the difference between the 2 populations. The greater the absolute value of SSMD, the larger the magnitude of difference between the 2 populations. For example, SSMD $= 3$ means that the mean is 3 times the standard deviation of the difference. In addition, SSMD has a link, with the probability of the difference being positive, as presented below.

d^+-probability

The probability of the difference being positive is called positive difference probability (denoted by d^+-probability), namely, d^+-probability $= \Pr(D > 0)$. This probability is a function of SSMD, that is, d^+-probability $= \Pr(Z_D > -\beta) = 1 - \Pr(Z_D \leq -\beta) = 1 - F_Z(-\beta)$, where $Z_D = \frac{D - \mu_D}{\sigma_D}$ and $F_Z(\cdot)$ is the cumulative distribution function of the standardized difference Z_D. When D is symmetrically distributed, d^+-probability $= \Pr(Z_D < \beta)$. Hence, the value of SSMD reflects the probability of the difference being positive. The larger the SSMD value, the greater the probability of the difference being positive and the greater the possibility that a value from the 1st group is larger than a value from the 2nd group.

In the situation in which D is normally distributed, the relationship between SSMD and d^+-probability is as simple as d^+-probability $= \Phi(\beta)$. In the situation in which D has a symmetric unimodal (or a unimodal) distribution with finite variance, d^+-probability has a low bound when $\beta \geq 1$ and has an up bound when $\beta \leq -1$. The relationships between SSMD and d^+-probability are summarized in **Table 1** for the 3 situations in which D has a symmetric unimodal distribution with finite variance, a unimodal distribution with finite variance, and a normal distribution, respectively. For example, given any value of β with $\beta \geq \sqrt{\frac{8}{3}}$, its corresponding d^+-probability is at least $1 - \frac{4}{9\beta^2}$ when the difference has a unimodal distribution with finite variance, at least $1 - \frac{2}{9\beta^2}$ when the difference has a symmetric unimodal distribution with finite variance, and is exactly $\Phi(\beta)$ when the difference has a normal distribution. The relationships between SSMD and d^+-probability

Table 1. The Relationships Between the Strictly Standardized Mean Difference (SSMD) and d^+-Probability

	D Has a Unimodal Distribution with Finite Variance		*D Has a Symmetric Unimodal Distribution with Finite Variance*		*D Has a Normal Distribution*
Range of SSMD Value	*Low Bound of* d^+*-Probability*	*UP Bound of* d^+*-Probability*	*Low Bound of* d^+*-Probability*	*UP Bound of* d^+*-Probability*	*Value of* d^+*-Probability*
$\beta \geq \sqrt{\frac{8}{3}}$	$1-\frac{4}{9\beta^2}$	/	$1-\frac{2}{9\beta^2}$	/	$\Phi(\beta)$ where $\Phi(\cdot)$ is the cumulative distribution function of $N(0,1)$
$\sqrt{\frac{8}{3}} > \beta \geq 1$	$\frac{4}{3}-\frac{4}{3\beta^2}$	/	$\frac{7}{6}-\frac{2}{3\beta^2}$	/	
$1 > \beta > -1$	/	/	/	/	
$-\sqrt{\frac{8}{3}} < \beta \leq -1$	/	$\frac{4}{3\beta^2}-\frac{1}{3}$	/	$\frac{2}{3\beta^2}-\frac{1}{6}$	
$\beta \leq = \sqrt{\frac{8}{3}}$	/	$\frac{4}{9\beta^2}$	/	$\frac{2}{9\beta^2}$	

provide a basis for interpreting the magnitude of difference from a probability perspective.

Statistical estimation and inference of SSMD

The concept of SSMD is based on population level. In reality, the population value of SSMD is rarely known. We usually have samples on which we can make an estimation and inference about the unknown parameters. Suppose we have 1 sample of size n_1, namely, $X_{11}, X_{12}, \ldots, X_{1n_1}$, being independently identically distributed from population P_1 and another independent sample of size n_1, namely, $X_{21}, X_{22}, \ldots, X_{2n_2}$, being independently identically distributed from population P_2. Let $N = n_1 + n_2$, let $\bar{X}_1$ and s_1 be the sample mean and sample standard deviation from the 1st sample, and let $\bar{X}_2$and s_2 be the sample mean and sample standard deviation from the 2nd sample, respectively.

Zhang[19] derived the statistical estimation and inference of SSMD when the 2 compared groups have unequal variances. That is, the maximum likelihood estimate (MLE) of SSMD based on independent groups is $\frac{\hat{\beta} = \bar{X} - \bar{X}}{\sqrt{\frac{n_1-1}{n_1}s_1^2 + \frac{n_2-1}{n_2}s_2^2}}$, and the MLE of the variance of $\hat{\beta}$ is

$$\hat{\sigma}^2_{\hat{\beta}} = \frac{\frac{n_1-1}{n_1^2}s_1^2 + \frac{n_2-1}{n_2^2}s_2^2}{\frac{n_1-1}{n_1}s_1^2 + \frac{n_2-1}{n_2}s_2^2} + \left(\frac{(n_1-1)^2}{n_1^3}s_1^4 + \frac{(n_2-1)^2}{n_2^3}s_2^4\right)\frac{(\bar{X}_1-\bar{X}_2)^2}{2\left(\frac{n_1-1}{n_1}s_1^2 + \frac{n_2-1}{n_2}s_2^2\right)^3}.$$

Based on the asymptotical normality of MLE, the $1-\alpha$ confidence interval of β is $\hat{\beta} \pm Z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}}$. Z_α is defined such that $\Pr(Z \leq Z_\alpha) = 1-\alpha$, and Z is a standard normal distribution.

When the 2 compared groups have equal variances, we can derive a better estimate of SSMD, the uniformly minimal variance unbiased estimate (UMVUE). The UMVUE of SSMD is

$$\hat{\beta} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{2}{K}\left((n_1-1)s_1^2 + (n_2-1)s_2^2\right)}},$$

$$K = 2\cdot\left(\frac{\Gamma\left(\frac{n_1+n_2-2}{2}\right)}{\Gamma\left(\frac{n_1+n_2-3}{2}\right)}\right)^2 \approx n_1+n_2-3.5 \quad \text{when } n_1, n_2 \geq 2.$$

The MLE of the variance of $\hat{\beta}_{\text{umvue}}$ is

$$\hat{\sigma}^2_{\hat{\beta}_{\text{UMVUE}}} = \frac{(n_1+n_2)K}{2n_1n_2(n_1+n_2-4)} + \frac{K-(n_1+n_2-4)}{n_1+n_2-4} \cdot \frac{(\bar{X}_1-\bar{X}_2)^2}{\frac{2}{n_1+n_2}\left((n_1-1)s_1^2+(n_2-1)s_2^2\right)} \approx \frac{1}{2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right).$$

The $1-\alpha$ confidence interval of β is approximately $\hat{\beta}_{\text{umvue}} \pm Z_{\frac{\alpha}{2}}\hat{\sigma}_{\hat{\beta}_{\text{UMVUE}}}$.

In primary HTS experiments, $n_1 = 1$ for most investigated siRNAs. In this case, the UMVUE of SSMD is then $\hat{\beta}_{\text{UMVUE}} = \frac{X_{11}-\bar{X}_2}{\sqrt{\frac{2}{K}(n_2-1)s_2^2}}$, where $K = 2\cdot\left(\frac{\Gamma\left(\frac{n_2-1}{2}\right)}{\Gamma\left(\frac{n_2-2}{2}\right)}\right)^2 \approx n_2 - 2.5$. And the MLE of the variance of $\hat{\beta}_{\text{UMVUE}}$ is then

$$\hat{\sigma}^2_{\hat{\beta}_{\text{UMVUE}}} = \frac{(1+n_2)K}{2n_2(n_2-3)} + \frac{K-(n_2-3)}{n_2-3}\cdot\frac{1+n_2}{2(n_2-1)}\cdot\frac{(X_{11}-\bar{X}_2)^2}{s_2^2} \approx \frac{1}{2}\left(1+\frac{1}{n_2}\right).$$

A NEW METHOD WITH A BALANCED CONTROL OF ERROR RATES

z-score method and t*-test of testing mean difference*

Currently, the mean ± k SD method and its variants, such as the median ± k MAD method, are commonly used in hit selection in HTS assays. SD and MAD denote standard deviation and median absolute deviation, respectively. These methods are based on the *p*-value from either *z*-score method

or the *t*-test for testing no mean difference. The mean ± k SD method addresses the question of what would happen if an investigated siRNA truly comes from the negative reference population with no specific inhibition or activation effects. That is, under the null hypothesis H_0: $\mu_1 = \mu_2$, we have $z\text{-score} = \frac{X_{11} - \bar{X}_2}{s_2} \sim N(0,1)$, approximately in the assumption of equal variance (namely, $\sigma_1^2 = \sigma_2^2 = \sigma^2$) and the assumption that $\bar{X}_2$ and s_2^2 can well represent μ_2 and σ_2^2, respectively. The mean ± SD method relies on the *z*-score of the standard normal distribution, N(0,1), and is thus also called the *z*-score method. The mean ± k SD method is widely used in hit selection in primary HTS experiments because it is hardly feasible to estimate the variance σ_1^2 in a tested siRNA because there is no replicate for most investigated siRNAs in most primary screens. The value of k is chosen so that the false-positive rate can be controlled to be no more than a preset level. This preset false-positive rate is usually either 0.05 or 0.01. The value of k is usually set to be 3 so that we can control the false-positive rate to be less than 0.05 for any unimodal distribution with finite variance. This is based on the well-known 3σ rule.[29] When the data are normally distributed, the mean ± 3 SD method actually controls the false-positive rate to be even smaller, namely, less than 0.0027, which comes from $2 \times (1 - \Phi(3))$. It is worthwhile to mention that the above formula is an approximation when the sample size n_2 in the negative reference group is large. This approximation might be problematic when n_2 is too small. When n_2 is small, a more appropriate formula for the *z*-score method is $z\text{-score} = \frac{X_{11} - \bar{X}_2}{s_2\sqrt{1+\frac{1}{n_2}}} \sim N(0,1)$, approximately. This is because the variance of $X_{11} - \bar{X}_2$ is $(1 + \frac{1}{n_2})\sigma_2^2$, not σ_2^2 in the condition of $\sigma_1^2 = \sigma_2^2$. Equivalently, mean ± k SD should be changed to mean $\pm k\sqrt{1 + \frac{1}{n_2}}$ SD when n_2 is small.

In confirmatory HTS experiments, the *t*-test of testing mean difference is popularly used for selecting hits because there are usually replicates for every tested siRNA or compound. It is well known that when there are replicates, the *t*-test is better than the *z*-score method for testing no mean difference in 2 groups, especially when the sample size is small. When the 2 compared groups have normal distributions, under the null hypothesis of H_0: $\mu_1 = \mu_2$,

$$t-\text{statistic} \begin{cases} \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{1}{n_1}+\frac{1}{2}\right)\left(\frac{n_1-1}{N-2}s_1^2 + \frac{n_2-1}{N-2}s_2^2\right)}} \sim t(N-2), \text{ if equal variance;} \\ \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \sim t(v) \end{cases}$$

v is the integer part of $\frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2/(n_1-1) + \left(s_1^2/n_1\right)^2/(n_1-1)}$ according to the Satterwhite option. The *t*-test usually requires $n_1 \geq 2$ and $n_2 \geq 2$. In many primary HTS assays, there is no replicate for most tested siRNAs; thus, the *t*-test of testing mean difference is usually inappropriate for hit selection in primary HTS assays.

A new decision process

Both the *z*-score method and *t*-test of testing no mean difference control the false-positive rate, in which we conclude that $\mu_1 \neq \mu_2$, whereas, actually, $\mu_1 = \mu_2$. Both address the question of whether the means in the 2 compared groups are exactly the same. In reality, our primary focus is not whether the siRNAs or compounds have average inhibition/activation effects that are exactly the same as the negative reference. Instead, we are interested in the siRNAs or compounds with a desired effect (such as a large effect). Accordingly, there are 2 major concerns in the process of selecting the siRNAs with a large effect in RNAi HTS experiments. First, we do not want the siRNAs with large effects to be selected as nonhits, and second, we do not want the siRNAs with small effects to be selected as hits. In terms of statistics, we want to control the false-negative rate in which the siRNAs with large effects are not selected as hits as small as possible. Meanwhile, we also want to control the false-positive rate in which the siRNAs with small effects are selected as hits as small as possible. It is well known that it is hard to achieve both a low false-negative rate and a low false-positive rate at the same time. Therefore, we need to search a hit selection method that maintains a balanced control of both error rates.

Let us first investigate the false-positive rate for large effects based on SSMD. If we want to select siRNAs with large positive effects (namely, with a large positive magnitude of difference) represented by $\beta \geq c$, where c ($c > 0$) is a preset level of minimal magnitude of difference that we want to achieve, we should control the error rate of concluding $\beta < c$, whereas, actually, $\beta \geq c$. This rate is reflected by the probability Pr(conclude $\beta < c$ given $\beta \geq c$). If we want to select siRNAs with large negative effects represented by $\beta \leq -c$, we should control the error rate of concluding $\beta > -c$, whereas, actually, $\beta \leq -c$. These error rates are false-negative rates. For example, in the case of selecting hits with large positive effects, the siRNAs with $\beta \geq c$ are true positives; thus, the rate to be controlled, Pr(conclude $\beta < c$ given $\beta \geq c$), is a false-negative rate that is more similar to a type II error than a type I error in the hypothesis testing of no mean difference.

The error rates Pr(conclude $\beta < c$ given $\beta \geq c$) and Pr(conclude $\beta > -c$ given $\beta \leq -c$) depend on the decision rule that we build on statistical samples. Based on the estimate $\hat{\beta}$ and its estimated variance, we can construct the decision rules as follows. Define Z_α such that $Pr(Z \leq Z_\alpha) = 1 - \alpha$, where Z is a random variable having the standard normal distribution. The decision rule for selecting siRNAs with large positive effects is

$$\text{Decision Rule I: } \begin{cases} \text{conclude } \beta \geq c, & \text{if } \hat{\beta} \geq c - Z_\alpha \hat{\sigma}_{\hat{\beta}}; \\ \text{conclude } \beta < c, & \text{if } \hat{\beta} < c - z_\alpha \hat{\sigma}_{\hat{\beta}}. \end{cases}$$

The decision rule for selecting siRNAs with large negative effects is

$$\text{Decision Rule II: } \begin{cases} \text{conclude } \beta \le -c, & \text{if } \hat{\beta} \le -c + Z_\alpha \hat{\sigma}_{\hat{\beta}}; \\ \text{conclude } \beta > -c, & \text{if } \hat{\beta} > -c + Z_\alpha \hat{\sigma}_{\hat{\beta}}. \end{cases}$$

As shown in the appendix, when the sample size is large, the false-negative rates of decision rules 1 and 2 are $\Pr(Z < \frac{c-\beta}{\hat{\sigma}_{\hat{\beta}}} - z_\alpha \text{ given } \beta \ge c)$ and $\Pr(Z > \frac{-c-\beta}{\hat{\sigma}_{\hat{\beta}}} + Z_\alpha \text{ given } \beta \le -c)$, respectively. Both have an up limit of α. For convenience, following the statistical convention in defining type II error level, let us call the up limit of a false-negative rate the false-negative level (FNL). Then, decision rules 1 and 2 have an FNL of α.

For the control of the false-positive rate in RNAi HTS experiments, we should consider the fact that scientists usually are not concerned with the inclusion of siRNAs with a fairly large or even weak effect, although smaller than the preset level c of large effects especially in primary screens. Let us assume that we can tolerate the false positives with $\beta > c_2$ (or $\beta < -c_2$), where $0 \le c_2 \le c$ for selecting the siRNAs with large positive (or negative) effects. That is, when we want to select the siRNAs with large positive effects, the false-positive rate that we are really concerned about is Pr(conclude $\beta \ge c$ given $b \le c_2$), not Pr(conclude $\beta \ge c$ given $\beta \le c$), and when we want to select the siRNAs with large negative effects, the false-positive rate that we are really concerned about is Pr(conclude $\beta \le -c$ given $\beta \ge -c_2$), not Pr(conclude $\beta \le -c$ given $\beta \ge -c$). For convenience, let us call this type of false-positive rate the restricted false-positive rate. As shown in the appendix, the restricted false-positive rates for decision rules 1 and 2 are $\Pr(Z \ge \frac{c-\beta}{\hat{\sigma}_{\hat{\beta}}} - Z_\alpha \text{ given } \beta \le c_2)$ and $\Pr(Z \le \frac{-c-\beta}{\hat{\sigma}_{\hat{\beta}}} + Z_\alpha \text{ given } \beta \ge -c_2)$, respectively, and both have an up limit of $\Phi(Z_\alpha - \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}})$. Again, for convenience, following the statistical convention in defining type I error level, let us call the up limit of the restricted false-positive rate of a decision rule the restricted false-positive level (RFPL). Then, the up limit $\Phi(Z_\alpha - \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}})$ is the RFPL for decision rules 1 and 2.

Therefore, FNL and RFPL in either decision rule are related to each other through $\text{RFPL} = \Phi(Z_{\text{FNL}} - \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}})$, which indicates that the estimated variance $\hat{\sigma}^2_{\hat{\beta}}$ of SSMD estimate and the preset levels c and c_2 all have an impact on the relationship between FNL and RFPL. Given c and $\hat{\sigma}^2_{\hat{\beta}}$, the relationship between FNL and RFPL relies on the value of c_2; in other words, given an FNL, RFPL decreases as c_2 decreases. For example, given $\hat{\sigma}^2_{\hat{\beta}} = 0.707$ and $c = 3$, when $\text{FNL} = 0.05$, RFPL decreases from 0.95, 0.591, 0.0118, 0.029, to 0.005 as c_2 decreases from 3, 2, 1, 0.5, to 0 (**Fig. 1A**). When $c_2 = c$, $\text{RFPL} = 1 - \text{FNL}$ for any $\hat{\sigma}^2_{\hat{\beta}}$, which means that a low FNL leads to a high RFPL or vice-versa (as shown by the light blue lines in **Fig. 1**). As a result, we cannot achieve both a low FNL and a low RFPL simultaneously when $c_2 = c$. Also, when $\hat{\sigma}^2_{\hat{\beta}}$ is large, decision rules 1 and 2 rarely achieve a low FNL and a low RFPL simultaneously, even if c_2 reaches its smallest value 0, as shown in **Figure 1B**. However, when $\hat{\sigma}^2_{\hat{\beta}}$ is small, we can obtain both a low FNL and a low RFPL simultaneously. For example, when $c_2 = 0.5$, we can obtain a low FNL of 0.039 and a low RFPL of 0.039 simultaneously in the case of $\hat{\sigma}^2_{\hat{\beta}} = 0.71$ and $c = 3$. Therefore, this new hit selection process may provide us with a flexible and balanced control of false-negative and false-positive rates depending on the experimental need.

How to use the new method for hit selection in primary HTS assays

The major objective of hit selection in RNAi HTS assays is to select the siRNAs with a desired effect such as a large effect. To do so, we need to determine the size of siRNA effects. Based on SSMD and its probability interpretation, a cutoff criterion called the 1-2-3 rule has been proposed for the determination of siRNA effects, which is displayed in **Table 2**. In this rule, the most important 3 thresholds are 1, 2, and 3. The SSMD values of 1, 2, and 3 have clear meanings: The size of the mean difference is 1, 2, and 3 times that of the standard deviation of the difference. The d^+-probability associated with the SSMD of an siRNA is the probability that a value from this siRNA is greater than a value from the negative reference group. Based on **Table 1**, we can calculate the values or bounds of d^+-probability corresponding to the SSMD values of 1, 2, and 3. The d^+-probability values or bounds corresponding to the SSMD values of 1, 2, 3, and others are displayed in **Table 3**. From **Table 3**, SSMD = 1, 2, and 3 (or −1, −2, and −3) also indicates that the low bounds (or up bounds) of the corresponding d^+-probability are about 0.5, 0.95, and 0.975 (or 0.5, 0.05, 0.025), respectively, in the situation in which the difference has a symmetric unimodal distribution with finite variance. The probability values of 0.95 and 0.975 (or 0.05 and 0.025) are all commonly used to indicate large (or small) chances for an event.

The 1-2-3 rule can help us to classify the effect size based on the magnitude of difference. However, this rule is based on the population value of SSMD. In practice, the population value of SSMD is usually unknown. We may use the estimated SSMD value to approximately represent the population value when the sample size is large. However, in many cases, the sample size is small in real RNAi HTS experiments. In this

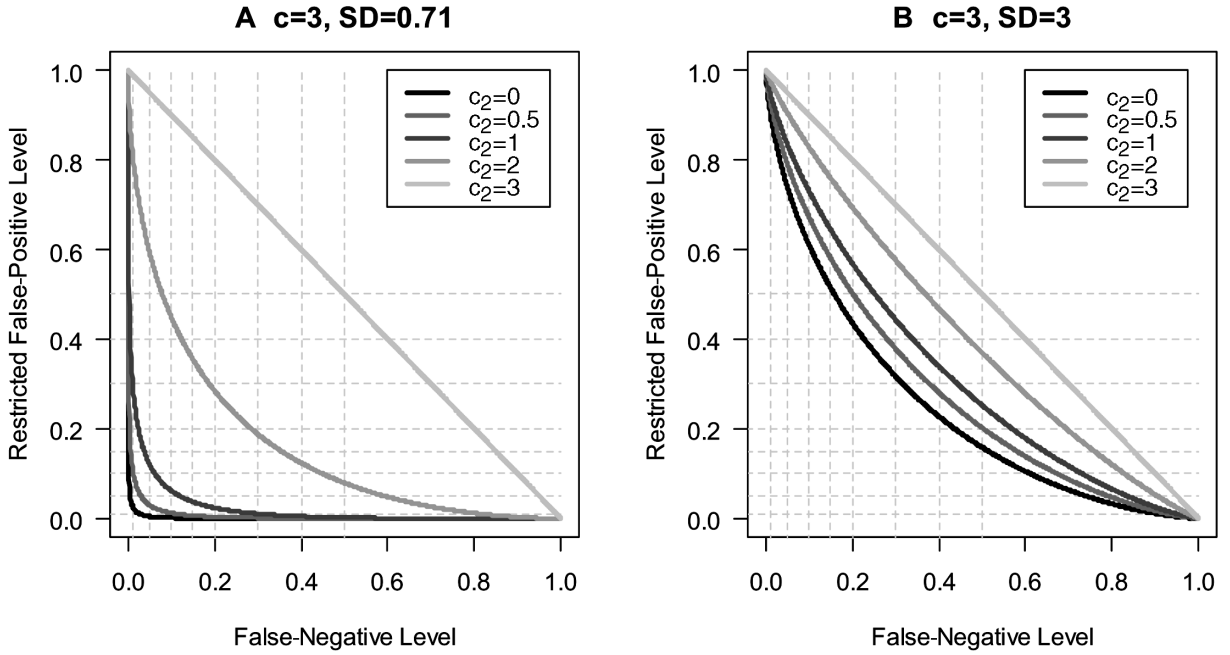

FIG. 1. Displaying the relationship between false-negative level and restricted false-positive level for 5 values (0, 0.5, 1, 2, and 3) of c_2 in the cases of $\hat{\sigma}^2_{\hat{\beta}} = 0.71$, c = 3 (**A**) and c = 3, c = 3 (**B**).

situation, we may need to make an inference of SSMD based on statistical samples and control the false-negative and false-positive rate to select the siRNAs with the desired effect sizes.

In RNAi HTS experiments, especially in the primary screens, we want to control a low false-negative rate in which the siRNAs with large effects (namely, those labeled as *strong*, *very strong*, or *extremely strong* in **Table 2**) are not selected as hits. Meanwhile, we want to maintain a low restricted false-positive rate in which the siRNAs with very weak or extremely weak effects are selected as hits. The balanced control of these 2 error rates can be achieved by applying the 2 SSMD-based decision rules in the case of $c = 3$, $c_2 = 0.5$, and 1. To do so, we need to calculate the estimated SSMD value $\hat{\beta}$ and determine the threshold $c - Z_{\mathrm{FNL}}\hat{\sigma}_{\hat{\beta}}$ in decision rule 1 or $-c + Z_{\mathrm{FNL}}\hat{\sigma}_{\hat{\beta}}$ in decision rule 2. For simplicity, let us focus on how to use decision rule 1 for selecting hits with a large positive siRNA effect in primary RNAi HTS experiments.

To calculate the estimated SSMD value and its estimated variance, we need to specify a negative reference group (the 2nd group in the formula for calculating SSMD) to represent siRNAs with no specific effects. In a typical siRNA HTS experiment, a negative control (such as luciferase) is designed to represent siRNAs with no specific effects. Thus, we may use this negative control to calculate SSMD. However, in a fairly large number of HTS experiments, the negative control is misleading because of position effects or assay variability.[15,16] As a result, it is common to use the sample wells instead of negative controls as a negative reference to calculate the percentage inhibition and *z*-score in the primary screen.[15–18,22–24] With this consideration, we construct the negative reference as follows: Exclude the largest 1% values and the smallest 1% values, and use the remaining 98% values in the sample wells in a plate as a negative reference group. The estimated SSMD between each investigated siRNA and this negative reference should then be used to detect the effect of this siRNA. There are about 100 to 320 investigated siRNAs, each with no replicate, in a 16 × 24 plate of a typical primary screen. That is, $n_1 = 1$ and n_2 ranges from 100 to 320; accordingly $\hat{\sigma}_{\hat{\beta}} \approx \sqrt{\frac{1}{2}(1 + \frac{1}{n_2})} \approx \sqrt{\frac{1}{2}} \approx 0.71$. With $\hat{\sigma}_{\hat{\beta}} \approx 0.71$ and $c = 3$, the threshold is $3 - 0.71 \times Z_{\mathrm{FNL}}$, and the relationship between FNL and RFPL is then $\mathrm{RFPL} = \Phi(Z_{\mathrm{FNL}} - \frac{3 - c_2}{0.71})$. **Figure 2** displays this relationship with $c_2 = 0$, 0.5, 1, 1.645, and 2, respectively.

There are 3 main approaches to determine the threshold of the estimated SSMD value $\hat{\beta}$ for hit selection. The 1st approach focuses on the control of a low FNL for SSMD ≥ c (and/or a low RFPL for SSMD ≤ c_2) depending on different experimental needs. In this approach, we may start with some common FNLs such as FNL = 0.01, 0.025, 0.05, and 0.10 and then simultaneously consider the RFPLs corresponding to some common c_2 values to choose a low FNL so that a low RFPL corresponding to a relatively large c_2 value can also be

Table 2. The 1-2-3 Rule for the Use of the Strictly Standardized Mean Difference to Determine Short Interfering RNA Effects

Cutoff Criterion for Effect Types	*Effect Type*
$\beta \geq 3$	Strong positive effect
$3 > \beta \geq 2$	Fairly strong positive effect
$2 > \beta \geq 1$	Moderate positive effect
$1 > \beta > 0$	Weak positive effect
$\beta = 0$	No effect
$-1 < \beta < 0$	Weak negative effect
$-2 < \beta \leq -1$	Moderate negative effect
$-3 < \beta \leq -2$	Fairly strong negative effect
$\beta \leq -3$	Strong negative effect
$\beta \geq 6.67$	Extremely strong positive effect
$6.67 > \beta \geq 4.7$	Very strong positive effect
$4.7 > \beta \geq 3$	Strong positive effect
$3 > \beta \geq 2$	Fairly strong positive effect
$2 > \beta \geq 1$	Moderate positive effect
$1 > \beta \geq 0.5$	Weak positive effect
$0.5 > \beta \geq 0.25$	Very weak positive effect
$0.25 > \beta > 0$	Extremely weak positive effect
$-0.25 > \beta < 0$	Extremely weak negative effect
$-0.5 < \beta \leq -0.25$	Very weak negative effect
$-1 < \beta \leq -0.5$	Weak negative effect
$-2 < \beta < -1$	Moderate negative effect
$-3 < \beta \leq -2$	Fairly strong negative effect
$-4.7 < \beta \leq -3$	Strong negative effect
$-6.67 < \beta \leq -4.7$	Very strong negative effect
$\beta \leq -6.67$	Extremely strong negative effect

achieved. **Table 4** lists 5 potentially used FNLs and their corresponding thresholds of $\hat{\beta}$ for hit selection.

For example, in some experiments, we may need a low FNL in which the siRNAs with large effects (namely, those labeled as *strong effect* in **Table 2**) are not selected as hits. In such cases, we may set FNL = 0.025 and the resulting decision rule, selecting an siRNA as a hit if it has $\hat{\beta} \geq 3 - 0.71 \times Z_{0.025}$ (namely, $\hat{\beta} \geq 1.6084$) and as a nonhit if $\hat{\beta} < 1.6084$. In this selection process, the RFPLs in which we select siRNAs with SSMD ≤ 0, SSMD ≤ 0.25 and SSMD ≤ 0.5 as hits are 0.0117, 0.0279, and 0.0592, respectively. These 3 RFPLs may be acceptable or marginally acceptable. If we set FNL = 0.01, the threshold is 1.3483, and the RFPL in which the siRNAs with SSMD ≤ 0.5 are selected as hits is 0.1161, which seems too high. Thus, FNL = 0.01 may not be a good choice unless we can tolerate many false positives with weak or no positive effects. In other experiments, we may need a low RFPL in which we include siRNAs with fairly small effects as hits. For example, we may need to control RFPL = 0.05 for SSMD ≤ 1 (namely, for weak or no positive effects). By solving $0.05 = \Phi(Z_{\text{FNL}} - \frac{3-1}{0.71})$, we obtain FNL = $1 - \Phi(\frac{3-1}{0.71} - 1.645) = 0.12$. Thus, the corresponding hit selection process is to select an siRNA as a hit if it has $\hat{\beta} \geq 3 - 0.71 \times Z_{0.12}$ (namely, $\hat{\beta} \geq 2.166$) and as a nonhit if $\hat{\beta} < 2.166$. Thus, the choice of FNL and RFPL and the consideration of c_2 are determined based on the need and objective of specific experiments.

The 2nd approach focuses on the balanced control of both FNL and RFPL. One case is to specify a c value and a c_2 value first and then to find the FNL value such that FNL = RFPL. This is a completely balanced control of 2 error levels, FNL and RFPL. For convenience, let us call this error level the completely balanced error level (CBEL). The CBELs corresponding to c = 3 and $c_2 = 0$, 0.5, 1, 1.645, and 2, respectively, are marked using bold numbers in **Figure 2**. Given c and c_2, the threshold of $\hat{\beta}$ is $\frac{c+c_2}{2}$ under the assumption of equal variance, and the corresponding error rate is $1 - \Phi(\frac{3-c_2}{0.71 \times 2})$. For example, the completely balanced error rate for c = 3 and $c_2 = 1$ is FNL = RFPL = $1 - \Phi(\frac{3-1}{1.42}) = 0.08$. Accordingly, the threshold is 2. In other words, an siRNA is selected as a hit if it has $\hat{\beta} \geq 2$ and as a nonhit if $\hat{\beta} < 2$. Under the normality condition, when the negative reference group works effectively, this process can ensure that both the false-negative rate in which the siRNAs with SSMD ≥ 3 (namely, strong positive effects) are not selected as hits and the restricted false-positive rate in which the siRNAs with SSMD ≤ 1 (namely, weak or no positive effects) are selected as hits are no more than 0.08.

The 3rd approach starts with the number of siRNAs that we want to select for further research (such as for confirmatory screen) and then calculates the corresponding FNLs and RFPLs. Suppose we want to select N_h siRNAs with large positive effects. We find the SSMD value c_h so that the number of siRNAs with $\hat{\beta} \geq c_h$ is N_h. Then solve $c_h = 3 - 0.71 \times Z_{\text{FNL}}$ to get FNL = $1 - \Phi(\frac{3-c_h}{0.71})$ and use RFPL = $\Phi(Z_{\text{FNL}} - \frac{3-c_2}{0.71})$ to obtain the RFPLs corresponding to some commonly used c_2 values.

We have presented the methods for selecting hits with large positive effects. Similarly, we can apply the methods for selecting hits with large negative effects.

DISCUSSION

In this article, we propose an SSMD-based method for hit selection. SSMD can assess the siRNA effect represented by the magnitude of difference between an siRNA and a negative reference group. The links between SSMD and d^+-probability offer us a clear interpretation of siRNA effects from a probability perspective (**Table 1**). The SSMD-based method of hit selection allows us to specify the size of siRNA effects including a large effect size c and a small effect size c_2 and then to have flexible and balanced controls of both the false-negative rate, in which the siRNAs with strong effects (SSMD $\geq$ c) are

Table 3. Strictly Standardized Mean Difference (SSMD) Values and Their Corresponding d^+-Probability Values (and Bounds) for Positive or Negative Effects When the Difference Has a Normal Distribution (and a Symmetric Unimodal Distribution)

SSMD (c_2 or c)		*0*	*0.25*	*0.5*	*1*	*1.645*	*2*	*3*	*4.7*	*6.67*
Positive effects	Value of d^+-probability (normal distribution)	0.50	0.60	0.69	0.84	0.975	0.95	0.99865		
	Low bound of d^+-probability (symmetric unimodal distribution)	—	—	—	0.5	0.944	0.918	0.975	0.99	0.995
Negative effects	Value of d^+-probability (normal distribution)	0.50	0.40	0.31	0.16	0.025	0.05	0.00135		
	UP bound of d^+-probability (symmetric unimodal distribution)	—	—	—	0.5	0.056	0.082	0.025	0.01	0.005

not selected as hits, and the restricted false-positive rate, in which the siRNAs with weak or no effects (SSMD $\leq c_2$) are selected as hits. The flexibility in the control of false negatives and false positives allows scientists to select hits based on the needs and objectives in their specific experiments. More important, when the scientists choose an SSMD-based process of hit selection, the FNL and RFPL in this process are readily known. The capability of balanced control in the SSMD-based methods allows scientists to achieve both a low false-negative rate in which the siRNAs with strong effects are not selected as hits and a low false-positive rate in which the siRNAs with weak or no effects are selected as hits.

The SSMD-based methods directly address the size of siRNA effects represented by the strength of difference between an siRNA and a negative reference, whereas the classic *z*-score method and *t*-test of testing no mean difference address whether the mean of an siRNA is exactly the same as the mean of a negative reference. Therefore, when we use the *z*-score method and *t*-test of testing mean difference for hit selection, the focus is the false-positive rate (namely, *p*-value) in which we conclude that the mean of an siRNA is different from the mean of the negative reference when the 2 means are actually the same. In contrast, when we use the SSMD-based method for hit selection, we first specify the strength of the difference in which we are interested in the experiments and then control both the false-negative rate and the restricted false-positive rate. In addition, the SSMD-based method can readily control the false-negative rate in which the siRNAs with large effects are not selected as hits, whereas it is nontrivial for the classic *z*-score method and *t*-test of testing mean difference to control the false-negative rate in which we conclude no mean difference while the means are different. Clearly, the SSMD-based method addresses scientific questions and fills scientific needs in practical experiments better than the currently used *z*-score method and *t*-test for hit selection.

To use the SSMD-based method for hit selection in RNAi HTS experiments, we need to calculate the estimated SSMD value $\hat{\beta}$ and to determine the threshold $c - Z_{\mathrm{FNL}}\hat{\sigma}_{\hat{\beta}}$ in decision rule 1 or $-c + Z_{\mathrm{FNL}}\hat{\sigma}_{\hat{\beta}}$ in decision rule 2. The goal in the determination of the thresholds is to maintain flexible and balanced control of FNL and RFPL depending on the experimental needs. To determine the threshold of $\hat{\beta}$ for hit selection, we should consider the balanced control of FNL, RFPL, c, and c_2 together. The classification of siRNA effects in **Table 2** provides a basis for the choice of c and c_2.

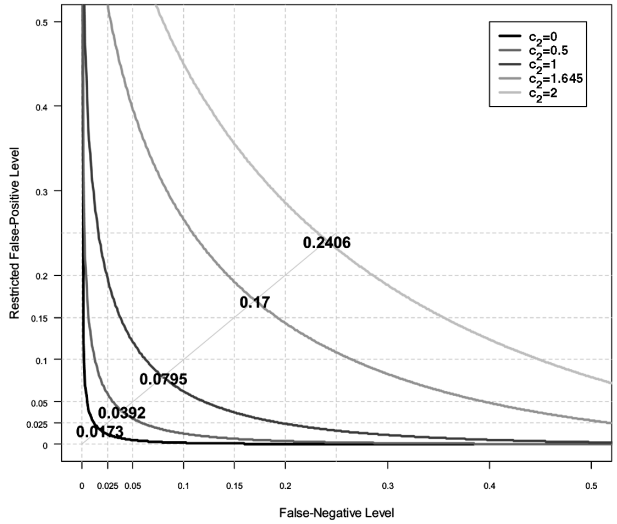

FIG. 2. The relationship between false-negative level and restricted false-positive level for selecting hits with large positive (or negative) effects in a typical primary RNA interference high-throughput screening experiment. A bold number on a curve denotes the completely balanced error level corresponding to a c_2 value ($c_2 = 0, 0.5, 1, 1.645, 2$).

There are 3 main approaches to determining the SSMD-based threshold for hit selection. The 1st approach focuses on the control of a low FNL for the siRNAs with strong effects and/or a low RFPL for the siRNAs with weak or no effects depending on different experimental needs. In this approach, we may start with some common FNLs listed in **Table 4** and then simultaneously consider the RFPLs corresponding to the $c_2$values that indicate weak or no effects such as 1, 0.5, 0.25, and 0. We may also start with weak effects represented by the

Table 4. Hit Selection Criteria and RFPLs for a Series of c_2 for 5 Potentially Used False-Negative Levels

		FNL = 0.01	*FNL = 0.025*	*FNL = 0.05*	*FNL = 0.10*	*FNL = 0.15*
Threshold of Estimated SSMD Value		$\hat{\beta} \geq 1.3483$	$\hat{\beta} \geq 1.6084$	$\hat{\beta} \geq 1.8322$	$\hat{\beta} \geq 2.0901$	$\hat{\beta} \geq 2.2641$
RFPL	A: $c_2 = 0$	0.0288	0.0117	0.0049	0.0016	0.0007
	B: $c_2 = 0.25$	0.0609	0.0279	0.0129	0.0048	0.0023
	C: $c_2 = 0.5$	0.1161	0.0592	0.0303	0.0126	0.0065
	D: $c_2 = 1$	0.3119	0.1957	0.1206	0.0623	0.0375
	E: $c_2 = 1.6449$	0.6619	0.5205	0.3960	0.2653	0.1916
	F: $c_2 = 2$	0.8207	0.7094	0.5934	0.4495	0.3549
	G: $c_2 = 3$	0.99	0.975	0.95	0.90	0.85

c_2 values and consider their corresponding FNLs and RFPLs. In the end, we determine a threshold so that we can maintain flexible and balanced control of both FNLs and RFPLs depending on the experimental needs. The 2nd approach is the CBEL approach, in which we specify a c value ($c = 3$) and a c_2 value first and then find the FNL value such that FNL = RFPL to obtain the threshold. The 3rd approach focuses on the number of siRNAs that we want to select for further research (such as for confirmatory screens). This approach appears to be similar to any other ranking method for hit selection. However, in the SSMD-based ranking method, for any given number of siRNAs with the largest effects, we know the up limits of the false-negative rate and restricted false-positive rates in the process of selecting these siRNAs as hits, which is a clear advantage of the SSMD ranking method over other ranking methods. In addition, when using the estimated SSMD value to approximately represent the population value, we are able to classify siRNAs according to the strength of their effects (**Table 2**).

As the SSMD-based methods of hit selection are novel methods currently under development, further research needs to be done. Like the classic *z*-score method and *t*-test of testing mean difference, the statistical estimation and inference of SSMD as well as the SSMD-based method of hit selection work ideally in the situation in which 2 compared groups are independent and each has a normal distribution. We need to investigate their robustness under nonnormality situations. Coupled with the robustness issue is sample size consideration. When the sample size is large, the asymptotical normality of SSMD estimates can ensure that the statistical estimation and inference of SSMD and SSMD-based methods for hit selection work well. For further research, we should investigate the statistical estimation and inference of SSMD as well as SSMD-based methods for hit selection when the sample size is small, especially when the sample size in the negative reference group is small. Although we used sample wells as the negative reference in this article, SSMD and SSMD-based methods should work if we use the negative control wells as the negative reference when the negative control wells work effectively. However, because of common issues with the negative control wells in RNAi HTS assays (such as strong bias and small number of negative control wells in a plate), we may need more research to explore the best strategy for constructing the negative reference group especially in confirmatory screens. Finally, the methods in this article have been developed from a statistical and methodological perspective based on scientific needs in RNAi HTS experiments. We need to demonstrate the practical usefulness of these methods in real RNAi HTS experiments. In a follow-up article,[28] we will report on the application of SSMD and SSMD-based methods in 2 in-house RNAi HTS experiments. Although the methods presented in this article are developed for hit selection in RNAi-based high-throughput screens, they should be applicable to other assays in which the end point is a difference in signal compared to a reference sample including those for receptor, enzyme, and cellular function.

ACKNOWLEDGMENT

The author thanks Drs. Daniel Holder, Keith Soper, and Joseph Heyse for their support in this research and the anonymous reviewers for their valuable comments.

APPENDIX

Derive false-negative rate and false-negative level

When the goal is to identify the short interfering RNAs (siRNAs) with large positive effects, the false negatives are the siRNAs that actually have strictly standardized mean difference (SSMD) $\geq$ c, but a decision rule or hit selection process leads us to conclude that they have SSMD $<$ c. The false-negative rate is thus Pr(conclude $\beta < c$ given $\beta \geq c$). Corresponding to decision rule 1 for selecting siRNAs with large positive effects, namely, the selection process of concluding $\beta \geq$ c if $\hat{\beta} \geq c - Z_\alpha \hat{\sigma}_{\hat{\beta}}$ and concluding $\beta g < c$ if

$\hat{\beta} < c - z_\alpha \hat{\sigma}_{\hat{\beta}}$, the false-negative rate is

$$\begin{aligned}&\Pr(\text{conclude } \beta < c \text{ given } \beta \geq c \text{ in decision rule 1})\\&= \Pr(\hat{\beta} < c - z_\alpha \hat{\sigma}_{\hat{\beta}} \text{ given } \beta \geq c)\\&= \Pr\left(\frac{\hat{\beta}-\beta}{\hat{\sigma}_{\hat{\beta}}} < \frac{c-\beta}{\hat{\sigma}_{\hat{\beta}}} - z_\alpha \text{ given } \beta \geq c\right)\\&= \Pr\left(Z < \frac{c-\beta}{\hat{\sigma}_{\hat{\beta}}} - z_\alpha \text{ given } \beta \geq c\right).\end{aligned}$$

$\Pr(Z < \frac{c-\beta}{\hat{\sigma}_{\hat{\beta}}} - z_\alpha \text{ given } \beta \geq c) \leq \Pr(Z < -Z_\alpha)$ because given $\beta \geq c$, $\frac{c-\beta}{\hat{\sigma}_{\hat{\beta}}} - Z_\alpha \leq -Z_\alpha$. When the sample size is large, $Z = \frac{\hat{\beta}-\beta}{\hat{\sigma}_{\hat{\beta}}}$ is approximately normally distributed; thus, $\Pr(Z \leq -Z_\alpha) = \alpha$. Therefore, the false-negative rate Pr(conclude $\beta < c$ given $\beta \geq c$ in decision rule 1) $\leq \alpha$. In other words, the FNL is α for decision rule 1.

When the goal is to identify siRNAs with large negative effects, the false negatives are the siRNAs that actually have SSMD $\leq -c$ ($c \geq 0$), but a decision rule or hit selection process leads us to conclude that they have SSMD $> -c$. The false-negative rate is thus Pr(conclude $\beta > -c$ given $\beta \leq -c$). Corresponding to decision rule 2 for selecting siRNAs with large negative effects, namely, the selection process of concluding $\beta \leq -c$ if $\hat{\beta} \leq -c + Z_\alpha \hat{\sigma}_{\hat{\beta}}$ and concluding $\beta > -c$ if $\hat{\beta} < -c + z_\alpha \hat{\sigma}_{\hat{\beta}}$, the false-negative rate is

$$\begin{aligned}&\Pr(\text{conclude } \beta < c \text{ given } \beta \leq -c \text{ in decision rule 2})\\&= \Pr(\hat{\beta} > -c + Z_\alpha \hat{\sigma}_{\hat{\beta}} \text{ given } \beta \leq -c)\\&= \Pr\left(\frac{\hat{\beta}-\beta}{\hat{\sigma}_{\hat{\beta}}} > \frac{-c-\beta}{\hat{\sigma}_{\hat{\beta}}} + Z_\alpha \text{ given } \beta \leq -c\right)\\&= \Pr\left(Z > \frac{-c-\beta}{\hat{\sigma}_{\hat{\beta}}} + Z_\alpha \text{ given } \beta \leq -c\right).\end{aligned}$$

$\Pr(Z > \frac{-c-\beta}{\hat{\sigma}_{\hat{\beta}}} + Z_\alpha \text{ given } \beta \leq -c) \leq \Pr(Z > Z_\alpha)$ because given $\beta \leq -c$, $\frac{-c-\beta}{\hat{\sigma}_{\hat{\beta}}} + Z_\alpha \geq Z_\alpha$. Again, when the sample size is large, $Z = \frac{\hat{\beta}-\beta}{\hat{\sigma}_{\hat{\beta}}}$ is approximately normally distributed; thus, $\Pr(Z > Z_\alpha) = \alpha$. Therefore, the false-negative rate Pr(conclude $\beta > -c$ given $\beta \leq -c$ in decision rule 2) is less than or equal to α. In other words, the FNL is α for decision rule 2.

Derive restricted false-positive rate and restricted false-positive level

When we want to identify siRNAs with large positive effects and also avoid including siRNAs with very weak or no positive effects in the hit list, the restricted false positives are the siRNAs that actually have SSMD $\leq c_2$, but a decision rule or hit selection process leads us to conclude that they have SSMD $\geq c$. The restricted false-positive rate is thus Pr(conclude $\beta \geq c$ given $\beta \leq c_2$). Corresponding to decision rule 1 for selecting siRNAs with large positive effects, the restricted false-positive rate is

$$\begin{aligned}&\Pr(\text{conclude } \beta \geq c \text{ given } \beta \leq c_2 \text{ in decision rule 1})\\&= \Pr\left(\hat{\beta} \geq c - Z_\alpha \hat{\sigma}_{\hat{\beta}} \text{ given } \beta \leq c_2\right)\\&= \Pr\left(\frac{\hat{\beta}-\beta}{\hat{\sigma}_{\hat{\beta}}} \geq \frac{c-\beta}{\hat{\sigma}_{\hat{\beta}}} - Z_\alpha \text{ given } \beta \leq c_2\right)\\&= \Pr\left(Z \geq \frac{c-\beta}{\hat{\sigma}_{\hat{\beta}}} - Z_\alpha \text{ given } \beta \leq c_2\right).\end{aligned}$$

$\Pr(Z \geq \frac{c-\beta}{\hat{\sigma}_{\hat{\beta}}} - Z_\alpha \text{ given } \beta \leq c_2) \leq \Pr(Z \geq \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}} - Z_\alpha)$ because given $\beta \leq c_2$, $\frac{c-\beta}{\hat{\sigma}_{\hat{\beta}}} - Z_\alpha \geq \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}} - Z_\alpha$. When the sample size is large, $\Pr(Z \geq \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}} - Z_\alpha) = \Phi(Z_\alpha - \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}})$. Therefore, the restricted false-positive rate Pr(conclude $\beta \geq c$ given $\beta \leq c_2$ in decision rule 1) is less than or equal to $\Phi(Z_\alpha - \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}})$. In other words, the RFPL is equal to $\Phi(Z_\alpha - \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}})$ for decision rule 1.

When we want to identify siRNAs with large negative effects and also avoid including siRNAs with very weak or no negative effects on the hit list, the restricted false positives are the siRNAs that actually have SSMD $\geq -c_2$, but a decision rule or hit selection process leads us to conclude that they have SSMD $\leq -c$. The restricted false-positive rate is thus Pr(conclude $\beta \leq -c$ given $\beta \geq -c_2$). Corresponding to decision rule 2 for selecting siRNAs with large negative effects, the restricted false-positive rate is

$$\begin{aligned}&\Pr(\text{conclude } \beta \leq -c \text{ given } \beta \geq -c_2 \text{ in decision rule 2})\\&= \Pr(\hat{\beta} \leq -c + Z_\alpha \hat{\sigma}_{\hat{\beta}} \text{ given } \beta \geq -c_2)\\&= \Pr\left(\frac{\hat{\beta}-\beta}{\hat{\sigma}_{\hat{\beta}}} \leq \frac{-c-\beta}{\hat{\sigma}_{\hat{\beta}}} + Z_\alpha \text{ given } \beta \geq -c_2\right)\\&= \Pr\left(Z \leq \frac{-c-\beta}{\hat{\sigma}_{\hat{\beta}}} + Z_\alpha \text{ given } \beta \geq -c_2\right).\end{aligned}$$

$\Pr(Z \le \frac{-c-\beta}{\hat{\sigma}_{\hat{\beta}}} + Z_\alpha \text{ given } \beta \ge -c_2) \le \Pr(Z \le \frac{-c+c_2}{\hat{\sigma}_{\hat{\beta}}} + Z_\alpha)$ because given $\beta \ge -c_2$, $\frac{c-\beta}{\hat{\sigma}_{\hat{\beta}}} - Z_\alpha \ge \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}} - Z_\alpha$. When the sample size is large, $\Pr(Z \le -\frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}} + Z_\alpha) = \Phi(Z_\alpha - \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}})$. Therefore, the restricted false-positive rate Pr(conclude $\beta \le -c$ given $\beta \ge -c_2$ in decision rule 2) is less than or equal to $\Phi(Z_\alpha - \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}})$. Namely, RFPL is equal to $\Phi(Z_\alpha - \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}})$ for decision rule 2.

In both decision rules, we have $\mathrm{FNL} = \alpha$ and $\mathrm{RFPL} = \Phi(Z_\alpha - \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}})$; thus, we have $\mathrm{RFPL} = \Phi(Z_{\mathrm{FNL}} - \frac{c-c_2}{\hat{\sigma}_{\hat{\beta}}})$.

REFERENCES

1. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC: Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 1998;391:806-811.
2. Hannon GJ, Zamore PD: Small RNAs, big biology: biochemical studies of RNA interference. In Hannon GJ (ed): *RNAi: A Guide to Gene Silencing*. New York: Cold Spring Harbor Laboratory Press, 2003:87-108.
3. Mahanthappa N: Translating RNA interference into therapies for human diseases. *Pharmacogenomics* 2005;6:879-883.
4. Check E: Silent running: the race to the clinic. *Nature* 2006;442: 614-615.
5. Zuck P, Murray EM, Stec E, Grobler JA, Simon AJ, Strulovici B, et al: A cell-based beta-lactamase reporter gene assay for the identification of inhibitors of hepatitis C virus replication. *Anal Biochem* 2004;334:344-355.
6. MacKeigan JP, Murphy LO, Blenis J: Sensitized RNAi screen of human kinases and phosphatases identifies new regulators of apoptosis and chemoresistance. *Nat Cell Biol* 2005;7:591-600.
7. Nybakken K, Vokes SA, Lin TY, McMahon AP, Perrimon NA: Genome-wide RNA interference screen in *Drosophila melanogaster* cells for new components of the Hh signalling pathway. *Nat Genet* 2005; 37:1323-1332.
8. Pelkmans L, Fava E, Grabner H, Hannus M, Habermann B, Krausz E, et al: Genome-wide analysis of human kinases in clathrin- and caveolae/raft-mediated endocytosis. *Nature* 2005;436:78-86.
9. Bard F, Casano L, Mallabiabarrena A, Wallace E, Saito K, Kitayama H, et al: Functional genomics reveals genes involved in protein secretion and Golgi organization. *Nature* 2006;439:604-607.
10. Espeseth AS, Huang Q, Gates AT, Xu M, Yu Y, Simon AJ, et al: A genome wide analysis of ubiquitin ligases in APP processing identifies a novel regulator of BACE1 mRNA levels. *Mol Cell Neurosci* 2006;33:227-235.
11. Zhang J-H, Chung TDY, Oldenburg KR: A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J Biomol Screen* 1999;4:67-73.
12. Zhang J-H, Chung TDY, Oldenburg KR: Confirmation of primary active substances from high throughput screening of chemical and biological populations: a statistical approach and practical considerations. *J Comb Chem* 2000;2:258-265.
13. Zhang J-H, Wu X, Sills MA: Probing the primary screening efficiency by multiple replicate testing: a quantitative analysis of hit confirmation and false screening results of a biochemical assay. *J Biomol Screen* 2005;10:695-704.
14. Wu X, Sills MA, Zhang J-H: Further comparison of primary hit identification by different assay technologies and effects of assay and screen measurement variability. *J Biomol Screen* 2005;10:581-589.
15. Gunter B, Brideau C, Pikounis B, Liaw A: Statistical and graphical methods for quality control determination of high-throughput screening data. *J Biomol Screen* 2003;8:624-633.
16. Brideau C, Gunter B, Pikounis B, Liaw A: Improved statistical methods for hit selection in high-throughput screening. *J Biomol Screen* 2003;8: 634-647.
17. Kevorkov D, Makarenkov V: Statistical analysis of systematic errors in high-throughput screening. *J Biomol Screen* 2005;10:557-567.
18. Malo N, Hanley JA, Cerquozzi S, Pelletier J, Nadon R: Statistical practice in high-throughput screening data analysis. *Nat Biotechnol* 2006; 24:167-175.
19. Zhang XD: A pair of new statistical parameters for quality control in RNA interference high throughput screening assays. *Genomics* 2007; 89:552-561.
20. Zhang XD, Espeseth AS, Chung N, Ferrer M: Evaluation of a novel metric for quality control in an RNA interference high throughput screening assay. *Biocomp* 2006:385-390.
21. Zhang XD, Espeseth AS, Chung N, Holder DJ, Ferrer M: The use of strictly standardized mean difference for quality control in RNA interference high throughput screening experiments. *American Statistical Association Proceedings* 2006:882-886.
22. Zhang XD, Yang XC, Chung N, Gates AT, Stec E, Kunapuli P, et al: Statistical analysis in RNA interference high throughput screening experiments. Paper presented at The First RNAi Meeting, Cold Spring Harbor, NY, November 2005.
23. Zhang XD, Yang XC, Chung N, Gates AT, Stec E, Kunapuli P, et al: Exploring statistical methods for hit selection in RNA interference high throughput screening experiments. *American Statistical Association Proceedings* 2005:723-729.
24. Zhang XD, Yang XC, Chung N, Gates AT, Stec E, Kunapuli P, et al: Robust statistical methods for hit selection in RNA interference high throughput screening experiments. *Pharmacogenomics* 2006;7:299-309.
25. Cohen J: The earth is round ($p < 0.05$). *Am Psychol* 1994;49:997-1003.
26. Kirk R: Practical significance: a concept whose time has come. *Educ Psychol Meas* 1996;56:746-759.
27. Tukey JW: *Exploratory Data Analysis*. Reading, MA: Addison-Wesley, 1977.
28. Zhang XD, Ferrer M, Espeseth AS, Marine S, Stec E, Crackower M, et al: The use of strictly standardized mean difference for hit selection in primary RNA interference high throughput screening experiments. *J Biomol Screen.* In press.
29. Vysochanskij DF, Petunin YI: Justification of the 3-sigma rule for unimodal distribution. *Theory of Probability and Mathematical Statistics* 1980;21:25-36.

Address correspondence to:
Xiaohua Douglas Zhang
Merck Research Laboratories
Biometrics Research, WP53B-120
PO Box 4
West Point, PA 19486

E-mail: xiaohua_zhang@merck.com