A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays

Ji-Hu Zhang, Thomas D. Y. Chung and Kevin R. Oldenburg

J Biomol Screen 1999; 4; 67
DOI: 10.1177/108705719900400206

JOURNAL OF BIOMOLECULAR SCREENING
Volume 4, Number 2, 1999

Original Articles

A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays

JI-HU ZHANG, THOMAS D. Y. CHUNG, and KEVIN R. OLDENBURG

ABSTRACT

The ability to identify active compounds ("hits") from large chemical libraries accurately and rapidly has been the ultimate goal in developing high-throughput screening (HTS) assays. The ability to identify hits from a particular HTS assay depends largely on the suitability or quality of the assay used in the screening. The criteria or parameters for evaluating the "suitability" of an HTS assay for hit identification are not well defined and hence it still remains difficult to compare the quality of assays directly. In this report, a screening window coefficient, called "Z-factor," is defined. This coefficient is reflective of both the assay signal dynamic range and the data variation associated with the signal measurements, and therefore is suitable for assay quality assessment. The Z-factor is a dimensionless, simple statistical characteristic for each HTS assay. The Z-factor provides a useful tool for comparison and evaluation of the quality of assays, and can be utilized in assay optimization and validation.

INTRODUCTION

RECENT ADVANCES IN DRUG TARGET IDENTIFICATION[1–3] and chemical compound library construction[4,5] necessitate corresponding advances in HTS.[6,7] Most assays routinely used in basic biochemical and cell biology research are not suitable for industrial-scale screening. This situation has lead to the development of a specialized discipline devoted to the design of assays that are optimized for speed, efficiency, signal detection, and low reagent consumption. Assays for HTS not only require small sample volume, high throughput, and robustness, but also require adequate sensitivity, reproducibility, and accuracy in order to discriminate among a very large number of compounds that span the entire range of activity. By using the current standard HTS methods, the majority of screening assays can be performed in reasonably high throughput via the use of automated liquid handling and signal detection systems.

In most HTS programs, each compound is tested only in singlet (or in duplicate). A high degree of accuracy and sensitivity in the assay is therefore critical for identifying active compounds (called "hits"). Due to the nature of each assay methodology and the perturbation introduced by instrumental and human-associated random error, all of the measurements from an assay contain a degree of variability, yet hits need to be identified in the presence of and despite such signal measurement variation. Intuitively, the lower the measurement variation, the higher the confidence that an identified hit is "real", i.e., that it will reconfirm upon retesting. Therefore, in the design and validation of HTS assays, an assessment of the screening data variability, by measures such as the standard deviation (SD) or coefficient of variation (CV), is critical in determining whether an assay can identify hits with confidence.

There is still not a generally accepted and systematic method that can be used to indicate or evaluate the "quality" of a HTS assay. A "high-quality" HTS assay must be able to identify, with a high degree of confidence, those few compounds that display significant biological activity. However, due to the lack of a systematic evaluation method, one frequently encounters "loosely" or even improperly used statistical terms in evaluation of HTS assays. HTS assays have both common and special features when compared to other statistical events. Recently, Sittampalam et al.[8] have reported the use of a "signal

DuPont Pharmaceuticals Research Laboratories, Leads Discovery, DuPont Pharmaceuticals Company, Wilmington, Delaware.

window" to evaluate the performance of HTS assays. The signal window concept, as described, is constructive and useful but still lacks uniformity and simplicity in analysis of the relevant parameters. This article introduces a much simpler statistical parameter to determine the suitability of an assay for HTS and demonstrates how to put it to practical use.

SCREENING ASSAY QUALITY vs. SIGNAL-TO-NOISE OR SIGNAL-TO-BACKGROUND RATIO

In a typical HTS assay, large numbers of compounds (usually in the range of 50,000 to 500,000 compounds) are tested. Hits are identified as the active compounds that either activate or inhibit the assay signal above a defined threshold value from the sample mean signal.

There are two expressions that have been used loosely to indicate the quality of an assay: signal-to-noise ratio (S/N) and signal-to-background ratio (S/B). Rigorously speaking, the S/N ratio is classically defined (from its original use in assessing radio signals) as:

$$S/N = \frac{\text{mean signal} - \text{mean background}}{\text{standard deviation of background}} \quad (1)$$

By this definition, the S/N ratio is only an indication of the degree of confidence with which a signal can be regarded as real, i.e., different from the associated background noise. The S/N ratio does not contain all the information needed in order to evaluate the quality of a screening assay (*vide infra*). A second expression is the signal-to-background ratio:

$$S/B = \frac{\text{mean signal}}{\text{mean background}} \quad (2)$$

This term is distinct from the S/N ratio with which it is often loosely interchanged. Because the S/B ratio does not contain any information regarding data variation, its inappropriateness in evaluation of an assay should be obvious.

The inherent problem with using either the S/N ratio or the S/B ratio is that neither of them takes into full account both the variability in the sample and background measurements and the signal dynamic range. Figure 1 illustrates a scenario where both S/N and S/B ratios fail to reflect the assay quality. Figure 1A displays an assay where the mean signal is 50 and the mean background is 10. Figure 1B displays an assay where the mean signal is 100 and the mean background is 10. In both cases the SD of the background is 3.3. In the case of Figure 1A, the S/N ratio is 12 and S/B ratio is 5, and in Figure 1B, the S/N ratio is 27 and the S/B ratio is 10. Comparison of the S/N or S/B ratios suggests that the assay as performed in Figure 1B has larger S/N and S/B ratios (the S/N ratio is 2.3-fold better and the S/B is 2-fold better than in Figure 1A). Empirically, however, it is obvious that the assay as run in Figure 1A would be more suitable for screening because the data varies less. It is also noteworthy that in Figure 1B, the absolute sample signal is twice as large as in Figure 1A, yet the assay as configured in Figure 1A is much more suitable for HTS. Therefore, although for some assay systems, under certain favorable conditions, the S/N or S/B ratio may be coincidentally in agreement with the HTS assay quality, neither expression alone actually reflects the assay quality. Clearly, a more convenient and systematic metric to assess and compare assay suitability is necessary, and this metric should incorporate the standard deviations (or CVs) of both the unknown samples and reference controls.

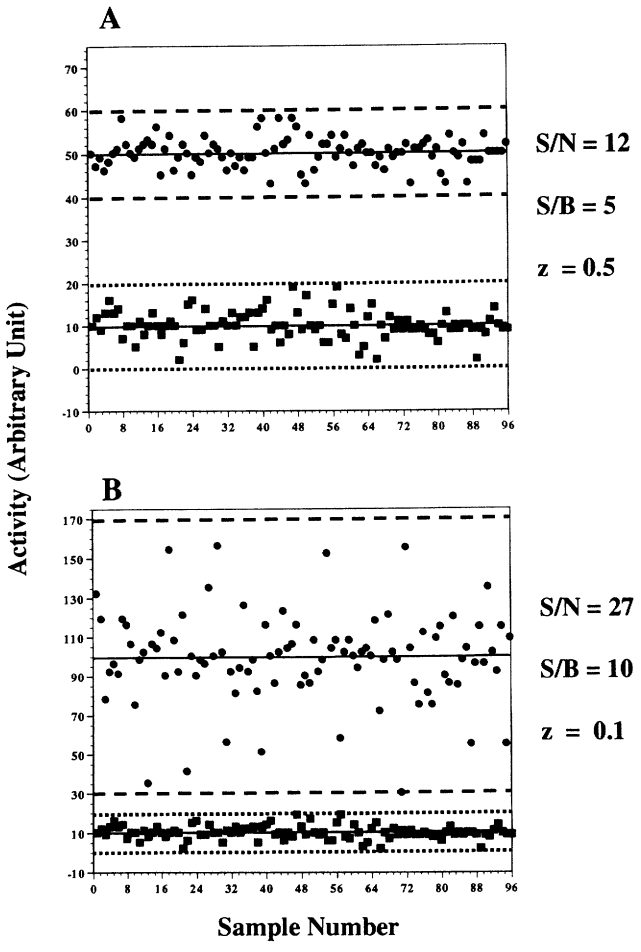

FIG. 1. Typical assay data from a single screen run in two different formats, **A** and **B**. The solid horizontal lines show the means of the sample and control (background) data in the two formats. Broken lines display 3 standard deviations (SD) from the mean of each data set. This data illustrates how the S/N or S/B ratios may be misleading in choosing the best format for the screen. The sample data in A has a mean of 50 with a 3 SD of 10, and in B the sample data has a mean of 100 with a 3 SD of 70. The control in both data sets has a mean of 10 with a 3 SD of 10. The S/N ratios are 12 and 27 and the S/B ratios are 5 and 10 for A and B respectively. By either S/N or S/B ratio, it would appear that the screen format that generated the data in B would be preferable for the screen. This error is due to the fact that neither the S/N nor the S/B ratio takes into account the variability in the sample data (which in many cases is different from that of the background). Empirically, and by calculation using the Z-factor, it is obvious that the screen format that generated the data in A is preferable for the screen.

SCREENING ASSAY QUALITY AND REPRODUCIBILITY OF HIT IDENTIFICATION

Rigorously, HTS assays should be analyzed using statistics that attempt to model the distribution of variables (biological responses) for a population of members (compounds). The distribution of a particular variable of a population may deviate significantly in modality, skewness, kurtosis, etc., from an ideal normal distribution. However, for large unbiased chemical libraries, the vast majority of the compounds have little or no biological activity. Therefore, error analysis by statistical models will apply, and an activity histogram of a typical HTS assay approximates a normal distribution model. The hit threshold or hit limit, is usually expressed as SD's away from the mean of the library sample signals. It is usually set at 3 SDs with a 99.73% confidence limit to the sample data. In practice, however, the "hit limit" is often set empirically so that the number of hits from the library can be handled reasonably well in secondary assays.

The hit limit not only affects the number of declared hits from a given screen, but also affects the number of false-positives and false-negatives. It is obvious that moving the hit limit farther from the sample mean lowers the false-positive rate for all the compounds with no activity and increases the false-negative rate for compounds with higher activities. However, due to the variation associated with each measurement, compounds with activity close to the hit limit always have some probability of crossing over the hit limit on remeasurement. Conceivably, compounds with an activity near the hit limit only have a certain probability of being declared a hit, as illustrated in Figure 2. Assuming the data variation for each compound also approximates the normal distribution profile, the probability of a compound being declared a hit and the probability of getting false-positive and false-negative hits can be assessed, as shown in Figure 3. For example, at the hit limit, approximately half (50%) of the compounds would be found with activities below the hit limit on retesting; at 3 SD further outside of the hit limit, a compound has approximately a 99.85% chance of being confirmed with an activity greater than the hit limit upon retesting. The number of false-positive and false-negative hits can only be minimized by reducing the data variability. This can be achieved either by assaying each data point multiple times, thus increasing the confidence of each data point, or by improving the assay quality. Thus, it can be concluded that, from a statistical point of view, the confirmation rate of the primary hits is affected by (1) the HTS assay quality, (2) the "hit limit" selection, and (3) the primary hit profile.

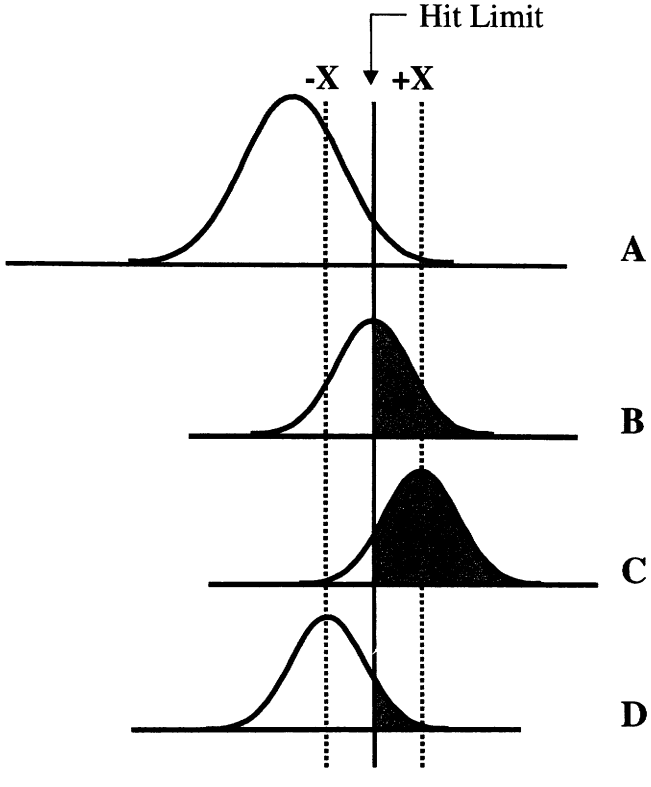

FIG. 2. A normal distribution of test samples (**A**) with a chosen "hit limit" (solid vertical line) and sample points (broken lines) taken at X on either side of the hit limit. Compounds with activity the same as the hit limit (**B**) will have an approximately 50% chance of being declared a hit (shaded area in B). Compounds with activities at (hit limit + X) or (hit limit − X) will have a much higher or lower probability of being declared a hit (**C** and **D**), respectively.

A STATISTICAL PARAMETER FOR HTS

In validating a typical HTS assay, unknown samples are assayed along with reference controls. The sample signal refers to the measured signal for a given test compound. The negative control (usually referred to as the background) refers to the set of individual assays from control wells that give the minimum signal. The positive control refers to the set of individual assays from control wells that give the maximum signal. In validating an assay, it is critical to run several assay plates containing both positive and negative controls in order to assess the reproduciblity and signal variation at the two extremes of the activity range. The positive and negative control data can then be used to calculate their respective means and SDs. The difference between the mean of the positive controls and the mean of the negative controls defines the dynamic range of the assay signal. The variation in signal measurement for samples, positive controls, and negative controls (i.e., SDs) may be different. The mean and SD of all the test sample signals are largely governed by the assay method and also by the intrinsic properties of the compound library. Because the vast majority of compounds from an unbiased library have very low or no biological activity, the mean and SD of all the sample signals should be close to those of the positive controls for inhibition/antagonist type assays and near those of the negative controls for activation/agonist type assays.

To define a screening window coefficient for HTS assays, μ_s, μ_{c-}, and μ_{c+} are denoted for the means of the library sample signal, negative control signal and positive control signal, respectively. The SDs of the signals are denoted as σ_s, σ_{c-}, and σ_{c+}, respectively. As discussed above, the difference of the means, $(\mu_{c+} - \mu_{c-})$, defines the assay dynamic range. Similar to the earlier described "noise band" definition,[8] the data vari-

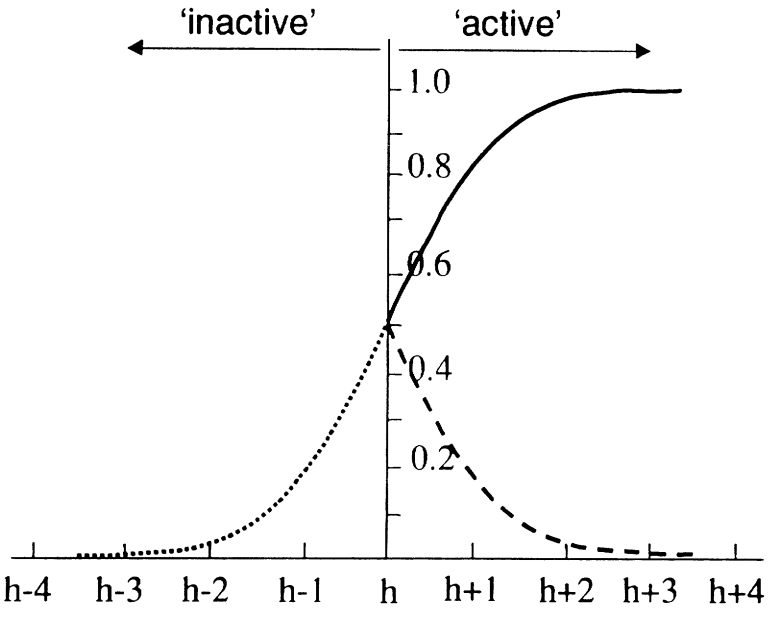

FIG. 3. The probability of correctly declaring a hit near the set "hit limit" of an HTS assay. In this diagram, the sample mean of the screened library lies somewhere to the left of the hit limit line (refer to Fig. 2). The horizontal axis displays activity away from the hit limit (h) in standard deviation (SD) units. The solid line (—) indicates the probability of a compound identified from the initial screen with activity greater than "h" being confirmed active on retest. The dashed line (----) indicates the probability of a compound being declared as having an activity above "h" in the primary assay when its "true" activity would be below "h" (false-positive), i.e., not confirmed on retest. The dotted line (·····) indicates the probability of a compound with an activity less than "h" in the initial screen having a "true" activity greater than "h" (false-negative). (Note that bias is built in because the compounds initially found below the hit limit are never rechecked because they are not scored.) Descriptive examples follow. A compound with an activity at "h + 3" in the primary assay would have a 99.85% chance of having an activity greater than "h." A compound with an initial activity at "h" would have only a 50% chance of having an activity greater than "h" on retest. A compound with initial activity of "h − 1" would have a 16% chance of having an activity greater than "h."

ation band (Fig. 4) for a given signal is defined as the mean ± 3 SDs, or $\mu \pm 3\sigma$ (i.e., 99.73% confidence limit). Thus hits are identified as those compounds whose signals are shifted away (usually by a defined SD) from the mean of the general sample population. The separation band between the sample and the control (Fig. 4), that is the useful window for identifying hits, is defined as $|\mu_s - \mu_c| - (3\sigma_c + 3\sigma_s)$, where $|\mu_s - \mu_c|$ defines the usable dynamic range for the screen, and is the absolute value of the difference of the two signal means. This expression for the separation band is suitable for either the activation type assay (where $\mu_c > \mu_s$) or the inhibition type assay (where $\mu_c < \mu_s$). Obviously, for the separation band, μ_c and σ_c need to be replaced by μ_{c+} and σ_{c+} for agonist/activation type assays and by μ_{c-} and σ_c for antagonist/inhibition type assays, respectively (Table 1).

It has already been shown that a signal window, is essential

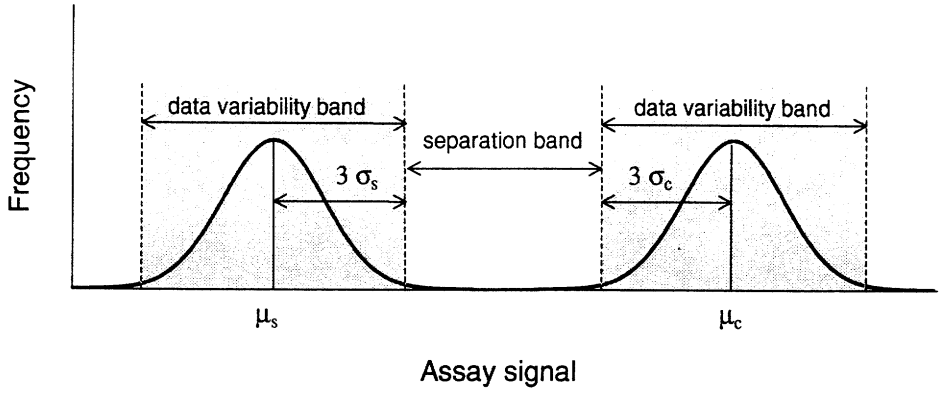

FIG. 4. Illustration of the defined data variation band and separation band in an HTS assay. The letter denotation is the same as described in the text. Both the sample and control data is assumed to obey a normal distribution profile. This should be a good approximation for large unbiased sample sets. For a small sample set, a t-distribution model may need to be considered.

TABLE 1. A SIMPLE CATEGORIZATION OF SCREENING ASSAY QUALITY BY THE VALUE OF THE Z-FACTOR

$$Z = 1 - \frac{3SD \text{ of sample} + 3SD \text{ of control}}{|\text{mean of sample} - \text{mean of control}|}*$$

Z-factor value	*Structure of assay*	*Related to screening*
1	SD = 0 (no variation), or the dynamic range $\rightarrow \infty$	An ideal assay
$1 > Z \geq 0.5$	Separation band is large	An excellent assay
$0.5 > Z > 0$	Separation band is small	A double assay
0	No separation band, the sample signal variation and control signal variation bands touch	A "yes/no" type assay
<0	No separation band, the sample signal variation and control signal variation bands overlap	Screening essentially impossible

*Note: For agonist/activation type assays the control data in the equation are substituted with the positive control (maximum activated signal) data; for antagonist/inhibition type assays the control data in the equation are substituted with the negative control (minimum signal) data. (see also Eq. 2 and Fig. 4 in the text.)

for hit identification in a HTS assay.[8] However, the use of the "signal window" (defined as the separation band) to evaluate assay quality lacks both clarity and uniformity. This is because the signal window can be expressed either in the same units as the directly measured signal or in the converted SD units of either the samples (σ_s) or the controls (σ_c). The selection of a proper SD unit for different types of assays (activation vs inhibition) in the signal window expression can be somewhat confusing and subjective. Therefore, a simple and dimensionless parameter is desirable for use in comparing and evaluating the quality of HTS assays. As such, a screening window coefficient (denoted Z-factor) is now defined as the ratio of the separation band to the signal dynamic range of the assay:

$$Z = \frac{|\mu_s - \mu_c| - (3\sigma_s + 3\sigma_c)}{|\mu_s - \mu_c|} \quad (3)$$

Or upon rearrangement,

$$Z = 1 - \frac{(3\sigma_s + 3\sigma_c)}{|\mu_s - \mu_c|} \quad (4)$$

This coefficient takes into account the assay signal dynamic range, the data variation associated with the sample measurement and the data variation associated with the reference control measurement. The Z-factor thus defines a characteristic parameter of the capability of hit identification for each given assay at the defined screening conditions (Table 1). It is therefore suitable to use the Z-factor for evaluating the quality (or performance) of HTS assays. In equation 4, the ratio $(3\sigma_s + 3\sigma_c)/|\mu_s - \mu_c|$ can be regarded as the data variation factor of the screening assay. It is clear that this variation factor by itself also takes into consideration all of the information necessary for HTS assay characterization.

The Z-factor is sensitive to the data variability as well as the signal dynamic range. For example, as $(3\sigma_s + 3\sigma_c)$ approaches zero, i.e., very small standard deviations, or as $|\mu_s - \mu_c|$ approaches infinity, the Z-factor approaches unity (maximum value of Z), and the HTS assay approaches an ideal assay. This is in agreement with one's intuition that a "good" assay has a large dynamic range and/or small data variability. The Z-factor can be any value less than or equal to one ($-\infty < Z \leq 1$). For any assay system, the Z value is only meaningful within the range of $-1 < Z \leq 1$. This is because at $Z = -1$, it is already at the lower detection limit of the assay system (*vide infra*). Because the Z-factor is a dimensionless quantity, it is suitable for comparison of assays, especially for an assay in different formats. The larger the value of the Z-factor of an assay, the higher the data quality (or the "suitability") of the assay for HTS (Fig. 1 and Table 1). Table 1 lists a simple comparison of screening assays by the value of their respective Z-factor.

The Z-factor can be used to evaluate the quality or performance of any given HTS assay. Similar to the Z-factor defined in equation 4 above, a Z′-factor can be calculated using only the control data:

$$Z' = 1 - \frac{(3\sigma_{c+} + 3\sigma_{c-})}{|\mu_{c+} - \mu_{c-}|} \quad (5)$$

The Z′-factor is a characteristic parameter for the quality of the assay itself, without intervention of test compounds. Therefore, the Z′-factor is a statistical characteristic of any given assay, not limited to HTS assays. Because in every case $Z \leq Z'$ for all large data sets with properly selected positive and negative reference controls, the Z′-factor can be utilized for quality assessment in assay development and optimization. For example, if the Z′ value is small (negative or close to zero), it usually indicates that the assay conditions have not been optimized or that the assay format is not feasible for generating useful data as it is configured. A comparison of the Z′ and Z values of the same assay under the same conditions reveals the effect of the compound library on the assay. If the Z′ value is large but the Z value is relatively small (where both Z′ and Z are >0), it may indicate that the compound library and/or the compound concentration for screening need to be further examined or optimized. Therefore, the Z′-factor is appropriate for evaluating overall assay quality, and the Z-factor reflects the quality of a configured assay for a particular HTS.

The Z-factor (Eq. 4), as with other statistical parameters, requires large data sets for improved accuracy. The reason lies with the high sensitivity of the Z-factor to data variability (σ's). Fluctuations in data variability will cause significant Z value changes, especially if the assay has a relatively narrow dynamic

range. This may explain why a large S/B ratio, although having very little meaning in assay quality evaluation, has been preferred in assay quality assessment in the past. The Z-factor expression clearly demonstrates that as the assay dynamic range shrinks, the Z-factor becomes more sensitive to changes in data variability. However, when the data variability is under tight control (low σ's), an assay with even a narrow dynamic range (thus a low S/B ratio) can still yield a large Z-factor and thus can be a high-quality assay.

In some reported cases, a 50% inhibition (midrange) reference control has also been used in assay quality control. In the Z-factor expression, the midpoint reference controls are not included for several reasons. First, the midpoint reference controls are not generally used throughout the industry. Not every assay has been carried out with a midpoint control, and the necessity of such controls is still the subject of debate. Secondly, the Z-factor is designed to be a simple and effective tool for assessing the quality of an assay, based on a few basic parameters such as the SDs and the dynamic range of an assay. Last, the Z-factor concept does not rule out but rather is compatible with such midpoint controls. It can be used to calculate a "half-range" Z value of any assay if the midpoint control data are substituted for the respective sample data.

Z-FACTOR IN HTS ASSAY OPTIMIZATION

As discussed, the Z-factor is an indicator of assay quality in HTS. It is conceivable that any change in assay conditions that affect the signal and signal variation will affect the Z value of the assay. Therefore the Z-factor is not only sensitive to the assay procedure and the instrumentation used, but also to the composition of the compound library and the compound concentration at which the screening is performed. Different libraries and different compound concentrations will give different distribution profiles of the sample signals. Consequently, the μ_s and σ_s values will change with these conditions, which causes a subsequent change of the Z-factor. Thus, the Z-factor can also be used as a tangible indicator in assay optimization. First, the assay should be optimized (as judged by the Z′-factor) for conditions (reagents, procedure, kinetics, instrument, etc.) other than those directly related to the test compounds. This ensures that the assay format has already been properly implemented and that the assay has sufficient dynamic range and acceptable signal variability and will provide useful data. Then, the properties of the compound library used for the screen should be considered. The suitable compound concentration for screening can be assessed by screening a representative subset of the library at several different concentrations (validation screening). From validation screening, the mean of the library sample, μ_s, standard deviation, σ_s, and subsequently the Z value, can be calculated at each compound concentration. In general, for a large unbiased library, the compound concentration that gives the largest Z value (at a desired hit rate) should be selected for the screen. One caveat of using the Z-factor to guide the optimization process is that it requires relatively large data sets. In practice, after the assay method is determined, several plates containing both positive and negative controls are assayed on at least 3 separate days in order to assess the well-to-well and day-to-day signal variation. A subset of plates (usually 20–40 plates) of compounds are then chosen to represent the entire library and tested at the defined screening conditions. The SDs (and CV%) and the assay dynamic range are then determined and used to calculate the Z-factor. If the Z-factor is sufficiently large (>0) at the defined conditions, then the assay can be used in HTS. Several HTS assays performed recently in our group have Z values in the range of 0.2–0.6.

It is important to note that the Z-factor as defined above (Eq. 4) is a plug-in formula useful for evaluating HTS assay quality. It has no preset requirement for setting the hit limit, and therefore, the hit limit is still a floating parameter of the screen. However, the Z value of an HTS assay will provide a useful guideline for where to set the most reasonable hit limit from the quality of the assay.

Z-FACTOR vs. CONFIRMATION RATE AND HIT QUALITY

A high-quality HTS assay should be able to identify hits with a high degree of confidence. The hit quality in this context refers to the confirmation rate from a primary screen. Because the Z-factor is characteristic of HTS assay quality, it is also related to the confirmation rate of the primary hits from the assay. However, the relationship between HTS assay quality (or the Z-factor value) and the hit confirmation rate for the assay is complicated by the compound library used in the screen. As discussed earlier, the hit confirmation rate (i.e., the percentage of confirmed positive hits out of the total hits collected from the primary screen) is a function of the assay quality, the hit limit selection, and the hit profile. The hit profile from a primary screen is largely governed by the compound library and compound concentration used in the screen. The difference in hit profiling will affect the confirmation probability of each hit (Fig. 3), and hence also affect the confirmation rate. Therefore, the Z-factor alone is only capable of predicting the relative confirmation rate when the same target is assayed by two or more different formats with the same compound library at the same compound concentration.

OTHER USES OF THE Z-FACTOR

The Z-factor concept can be utilized in areas other than assay development and HTS assay validation. For comparison of instrumentation quality, the two extreme reference controls over the dynamic range of the measurement can be used to calculate the Z′-factor, assuming that the data variation is approximately the interpolation of the two extreme controls. When this is true, the aforementioned data variation factor (i.e., $[3\sigma_{c+} + 3\sigma_{c-}]/|\mu_{c+} - \mu_{c-}|$ in Eq. 5) becomes a normalized, average data variability (with 99.73% confidence limit) of an instrument. For example, if Instrument A yields SDs of 0.02 and 0.04 at the two measurement extremes with a dynamic range of 0 to 10, then the Z′ value is 0.982. An improved version, Instrument B, yields SDs of 0.01 and 0.02 with the same dynamic range of 0 to 10. Then its Z′ value is 0.991. Consequently the data quality of Instrument A and B can be expressed by their respective Z′-factor values. An alternative approach is

to plot the Z values against each analyte concentration of a serial dilution over the entire range. This Z-C plot can be directly used for indicating and comparing the quality of a measuring device.

Another interesting link between the Z-factor and instrumentation is with the determination of the lower limit of detection (LLD), expressed as the mean of the background signal plus three SD of the background signal. When a measured signal is slightly larger than but approaching the background signal, i.e., $\mu_s > \mu_c$ and $\sigma_s \approx \sigma_c$, Eq. 4 yields: $\mu_s = \mu_c + 3\sigma_c$ at $Z = -1$. Therefore at $Z = -1$, the measured signal is equal to the lower limit of detection of the instrument.

SUMMARY

As discussed, it is suitable to utilize the Z-factor defined in this article in evaluation, comparison, and validation of any bioassays in general and of HTS assays in particular. For a real-world HTS assay, other aspects such as the reagent availability, the time line for the screen, and cost also need to be considered in the HTS assay design. Nevertheless, the ability to identify hits with high fidelity is the primary goal for HTS assay development. In summary, a screening window coefficient, the Z-factor, has been introduced in this report as a simple and tangible parameter to determine the suitability of an assay for HTS. It is a dimensionless characteristic for each HTS assay. In particular, the Z-factor can be utilized in HTS assay evaluation and validation.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Robert Nadon, Dr. Ilona Kariv, and Tongming Chen for helpful discussions and critical reading of the manuscript. A special thanks to Dr. Douglas Modlin for his input and suggestions on applications of the Z-factor to instrumentation.

REFERENCES

1. Cole, S.T., Saint Girons, I. (1994). Bacterial genomics. *FEMS. Microbiol. Rev.* 14: 2, 139–160.
2. Korenberg, J.R., Chen, X.N., Adams, M.D., Venter, J.C. (1995). Toward a cDNA map of the human genome. *Genomics* 29: 2, 364–370.
3. Fasman, K.H., Letovsky, S.I., Li, P., Cottingham, R.W., Kingsbury, D.T. (1997). The GDB Human Genetic Database Anno. 1997. *Nucleic Acids Res.* 25: 1, 72–81.
4. Gordon, E.M., Gallop, M.A., Campbell, D., Holmes, C., Bermak, J., Look, G., Murphy, M., Needels, M., Jacobs, J., Sugarman, J., Chinn, J., Ruhland-Fritsch, B., Jones, D. (1995). Combinatorial organic synthesis: Application to drug discovery. *Eur. J. Med. Chem.* 30:S337–S348.
5. Gordon, E.M., Barrett, R.W., Dower, W.J., Fodor, S.P.A., Gallop, M.A. (1994). Application of combinatorial technologies to drug discovery. 2. Combinatorial organic synthesis, library screening strategies and future directions. *J. Med. Chem.* 37:1385–1401.
6. Oldenburg, K.R., Zhang, J.-H., Chen, T., Maffia, A., Blom, K.A., Combs, A.P., Chung, T.D. (1998). Assay miniaturization for ultra-high throughput screening of combinatorial and discrete compound libraries: A 9600-well (0.2 microliter) assay system. *J. Biomol. Screen.* 3:55–62.
7. Stevens, M.E., Bouchard, P.J., Kariv, I., Chung, T.D-Y., Oldenburg, K.R. (1998). Comparison of automation equipment in high Throughput Screening. *J. Biomol. Screen*, 3:305–311.
8. Sittampalam, G.S., Iversen, P.W., Boadt, J.A., Kahl, S.D., Bright, S., Zock, J.M., Janzen, W.P., Lister, M.D. (1997). Design of signal windows in high throughput screening assays for drug discovery. *J. Biomol. Screen.* 2:159–169.

Address reprint requests to:
Dr. Ji-Hu Zhang
DuPont Pharmaceuticals Company
Experimental station, E400/5416
Rt. 141 and Henry Clay Rd.
Wilmington, DE 19880

E-mail: Ji-hu.Zhang@dupontpharma.com