

A tiered approach, using divergence from control as a measure of effect on one or more of these endpoints, categorizes doses into three levels no/low effect (L), medium effect (M), or high effect (H).

Effect size (ES) was defined using the variability in control wells:

$ES = \frac{S_{dose} - S_{ctrl}}{std_{ctrl}}$  where S is the statistic of interest,  $d$ ,  $P$ ,  $t_H$ , or  $t_P$ ,  $S_{ctrl}$  is the control data-estimated parameter,  $std_{ctrl}$  is the control-data estimated standard deviation of the parameter and represents the control noise/variability, and  $S_{dose}$  is defined as the lower 95% confidence limit, for  $S = P$  or  $S = t_H$ , or upper 95% confidence limit, for  $S = t_P$  or  $S = d$ , for that dose. For reliable estimates at least 3, and preferably 5 plates are needed.

Per Chebyshev's inequality (Chebyshev, 1867), 75% of the data is within 2 standard deviations of the mean, 94% is within 4 standard deviations of the mean and at least 96% is within 10 std of the mean. Using these criteria in the tiered approach:

- For parameters with greater values indicating greater dose effect, such as  $P$  or  $t_H$ 
  - If a dose's  $ES < 2$  then the dose is categorized as **no/low effect**.
  - If a dose's  $2 \leq ES \leq 4$  then the dose is categorized as **medium effect**.
  - If a dose's  $ES > 4$  then the dose is categorized as **high effect**.
- For parameters with lower values indicating greater dose effect, such as  $d$  or  $t_P$ 
  - If a dose's  $ES > -2$  then the dose is categorized as **no/low effect**.
  - If a dose's  $-4 \leq ES \leq 2$  then the dose is categorized as **medium effect**.
  - If a dose's  $ES < -4$  then the dose is categorized as **high effect**.

If the control wells show an early hyperactivity phase above baseline, this is likely due to plate swirling to distribute the dosing solutions, or slight temperature changes in the environment. Therefore, for  $t_H$ , it is unreasonable to use the control value as the "control  $t_H$ " because the control hyperactivity phase, if any, tends to be very short (1 hour or less). Thus, for  $t_H$ , it is preferable that the researcher specifies a reasonable value, e.g. 2 or 3 hours, as  $t_{H\_ctrl}$  to calculate the effect size and  $std_{ctrl}$  as 5% of that value, i.e. 0.1 (for 2 hours) or 0.15 (for 3 hours). The control standard deviation of 5% of the control value is based on the observed coefficient of variation (CV) of control values above baseline.