

Figure S1: Two probability density functions $P_{pos}(x)$ and $P_{neg}(x)$ representing positive and negative population distributions of an assay and a random variable x representing assay values. The risk of miss-classification with optimal threshold $\epsilon = \epsilon_0$ is represented by shaded area between two distribution overlap. where $R_1(h)$ represents False Negative Rate or type II error and $R_0(h)$ represents False Positive Rate or type I error.

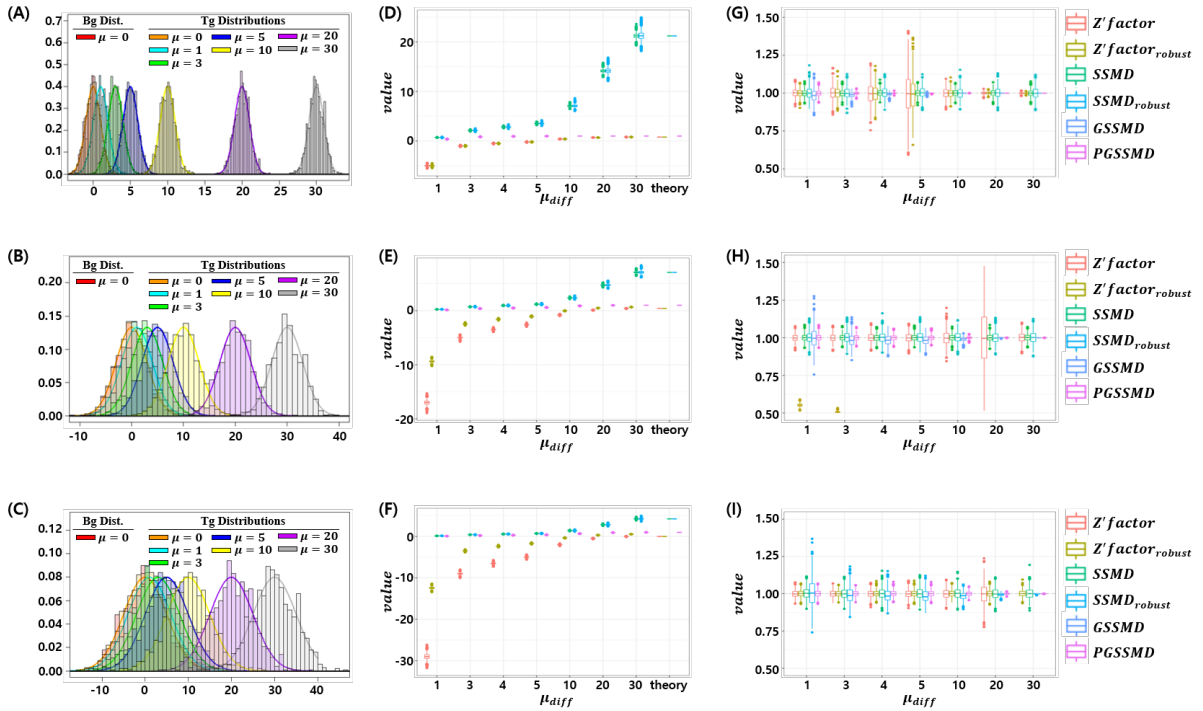


Figure S2: GSSMD provided interpretable and reliable statistical difference estimate in the ideal simulation setup ($\sigma_c = 1, 3, 5$, $n = 1,000$ in (A), (B), and (C) respectively). The results in (D) to (I) are acquired from random sampling of $N = 10^6$. (D), (E) and (F) are boxplots of each measure mean for a given mean difference and standard deviation. (G), (H), and (I) are boxplots of normalized measure for a given mean difference and standard deviation. Here we divided all values by corresponding theoretical value. The standard deviation of the sample SSMD was not sensitive to the mean difference, but the standard deviation of the sample GSSMD decreased as the mean difference increased.

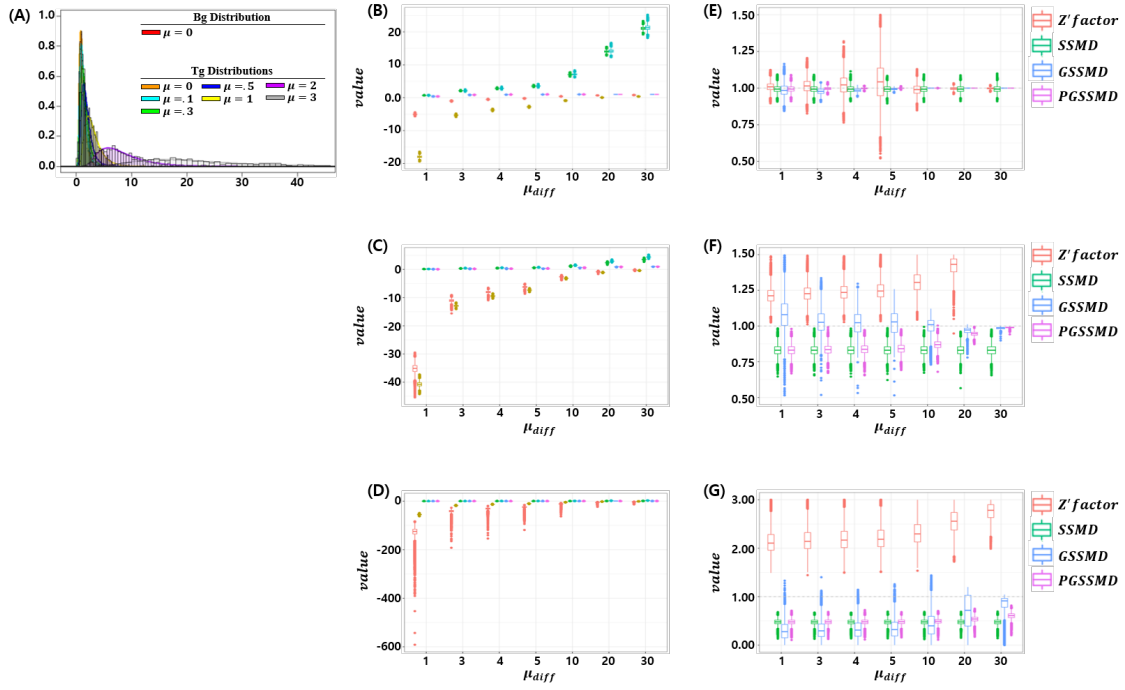


Figure S3: Performance of each measure in log-normal distribution setup. (A) shows log-normal distribution with $\sigma = 0.1$ and various mean differences. The results in (B) to (G) are acquired from random sampling of $N = 10^6$. (B), (C) and (D) are boxplots of each measure for a given mean difference and standard deviation. (E), (F), and (G) are boxplots of normalized (by its corresponding theoretical value) measure for a given mean difference and standard deviation.

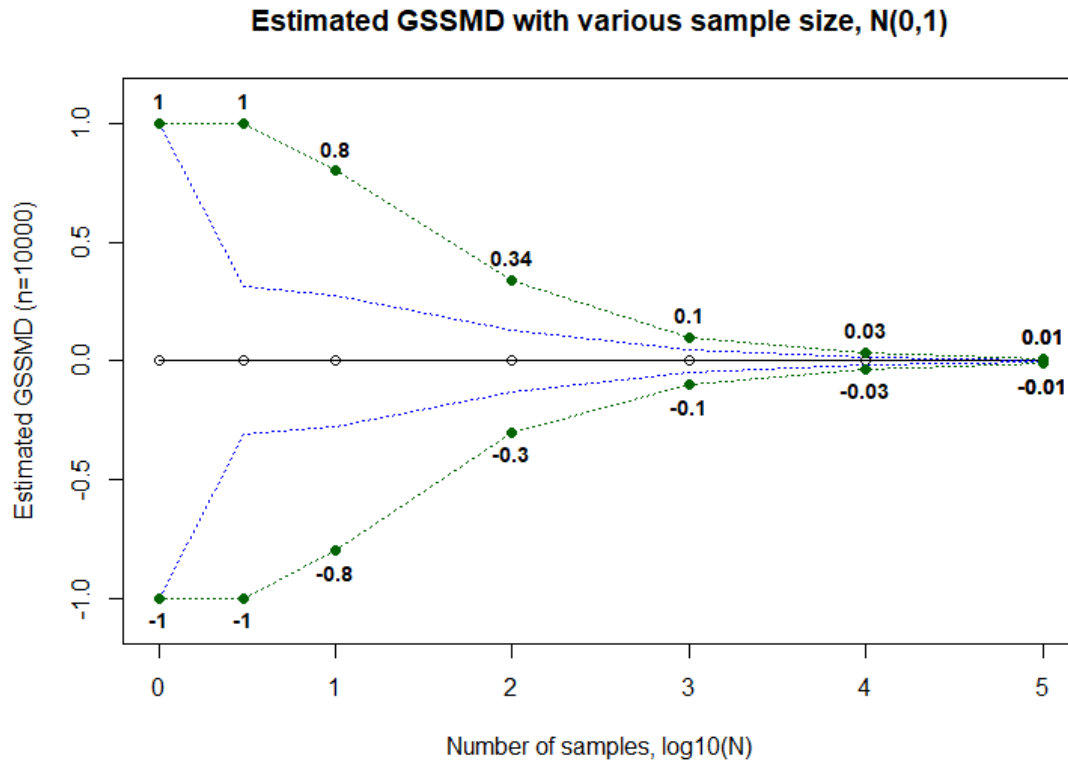


Figure S4: Effect of sample size on GSSMD when normal distributions overlap completely ($\mathcal{N}(0, 1)$). The blue line represents the standard deviation and the green line represents the maximum and minimum of the estimated GSSMD. The number of independent trials is $n = 10^4$

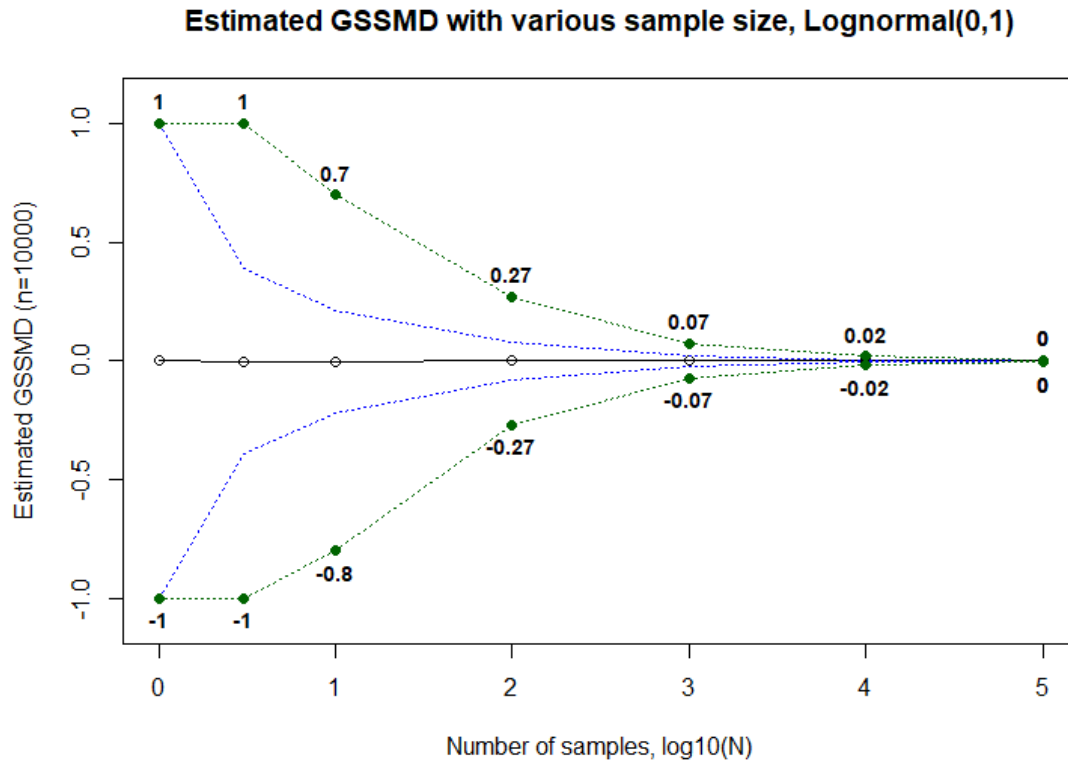


Figure S5: Effect of sample size on GSSMD when log-normal distributions overlap completely ($\text{Log}\mathcal{N}(0,1)$). The blue line represents the standard deviation and the green line represents the maximum and minimum of the estimated GSSMD. The number of independent trials is $n = 10^4$

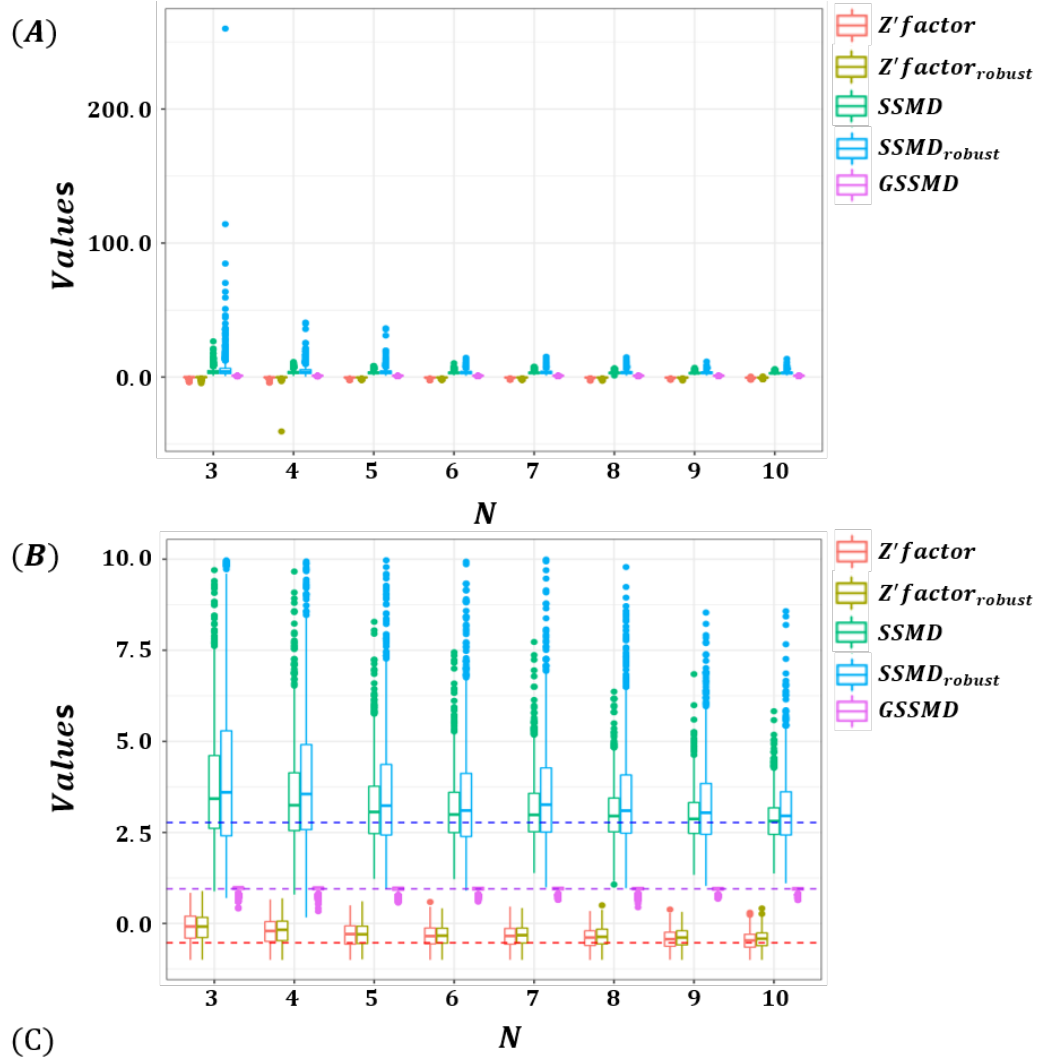


Figure S6: Parametric estimation of GSSMD showed least variance in the estimation of true effect size. $\mu_0 = 0, \mu_1 = 3.92, \sigma = 1$ for all sampling distributions. (A) Original scale plot showed large variance in SSMD and robust SSMD estimation. (B) y scale is modified to provide more detailed plot. (C) Calculated values of SSMD and GSSMD from the plot (A) and (B).

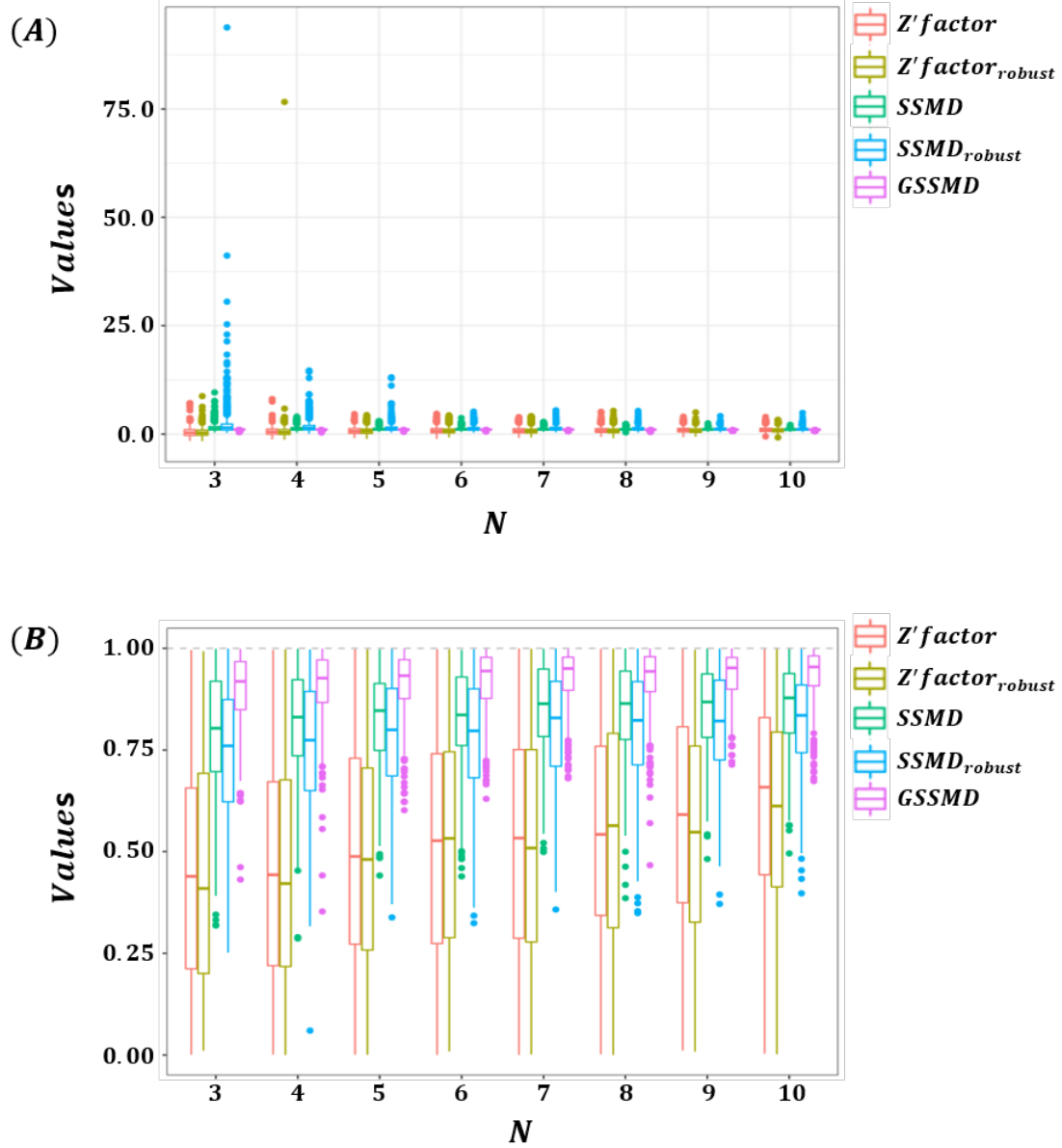


Figure S7: Parametric estimation of GSSMD showed least variance in the estimation of true effect size. (Scaled Plot) Here we scaled all values by its theoretical values. (A) Original scale plot showed large variance in SSMD and robust SSMD estimation. (B) y scale is modified to provide more detailed plot. In (B), We can see that GSSMD is the best estimator of true population effect size even though we have very small sample size ($N = 3 \rightarrow 10$). Therefore, we can use this parametric version of GSSMD for the very small sample size problem.

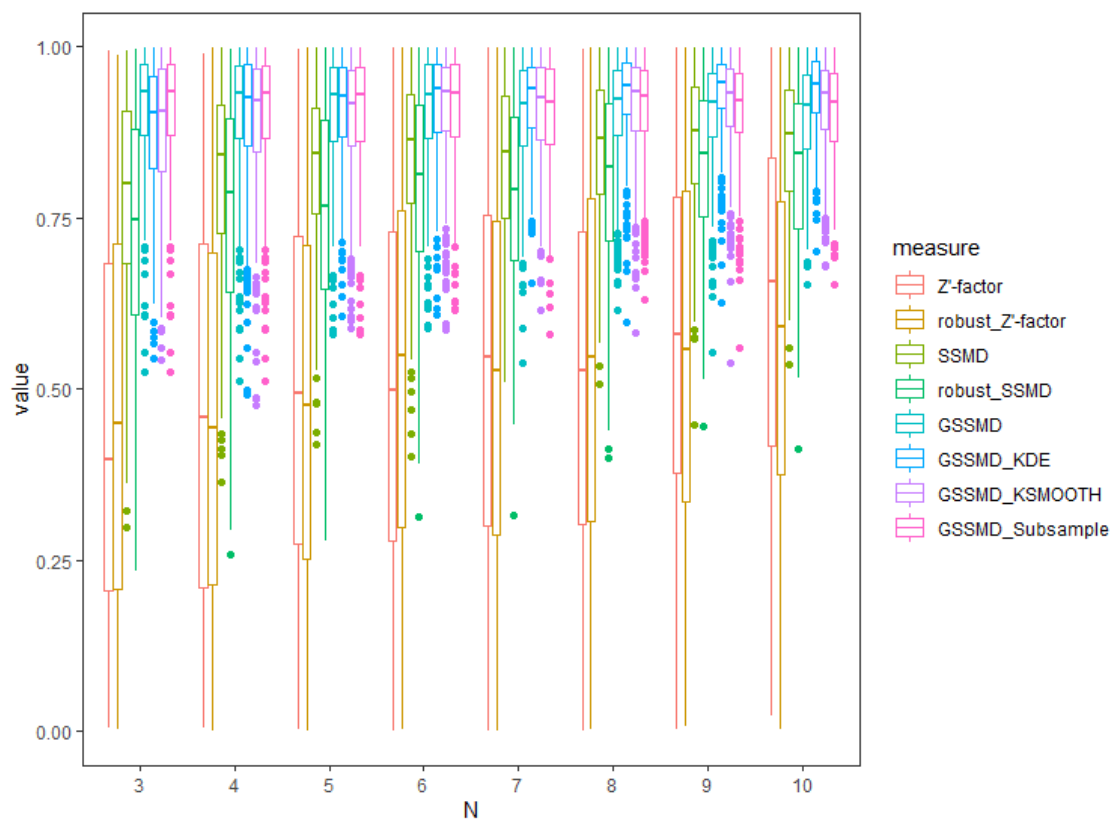


Figure S8: GSSMD are calculated using various methods such as histogram, KDE, KSMOOTH, and subsampling methods. These values are compared with Z'-factor and SSMD. The simulation setting is the same as figure S6 and S7.

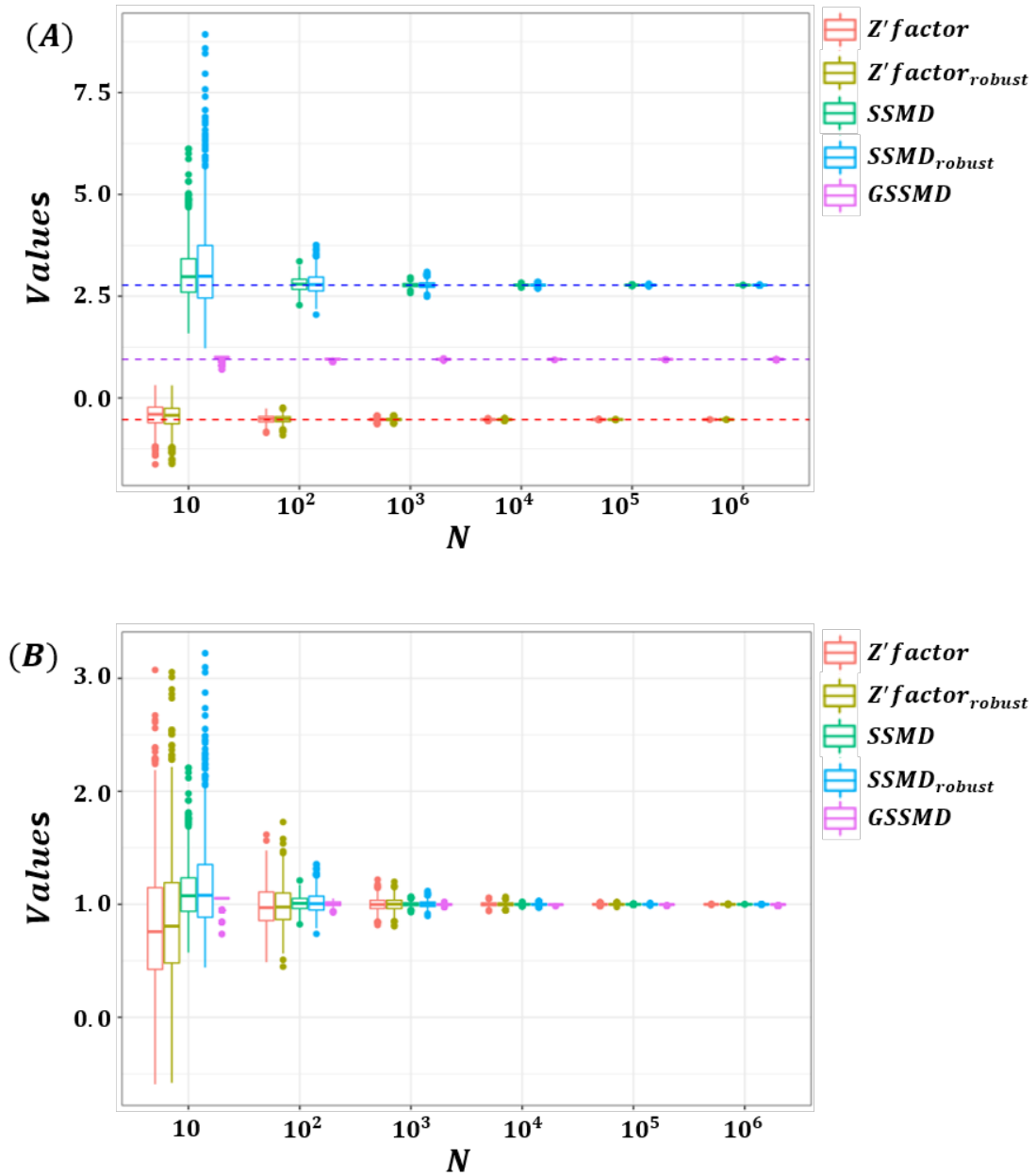


Figure S9: Estimation of GSSMD in large sample size. (A) GSSMD showed least variance in the estimation of true effect size where dashed lines illustrate true population effect size conditioned by the simulation. (B) We normalized each estimated measure by its theoretical value (SSMD = 2.77, GSSMD = 0.95, Z' -factor = 0.5). Similarly, as shown in the figure, GSSMD showed least variance compared to all measures.

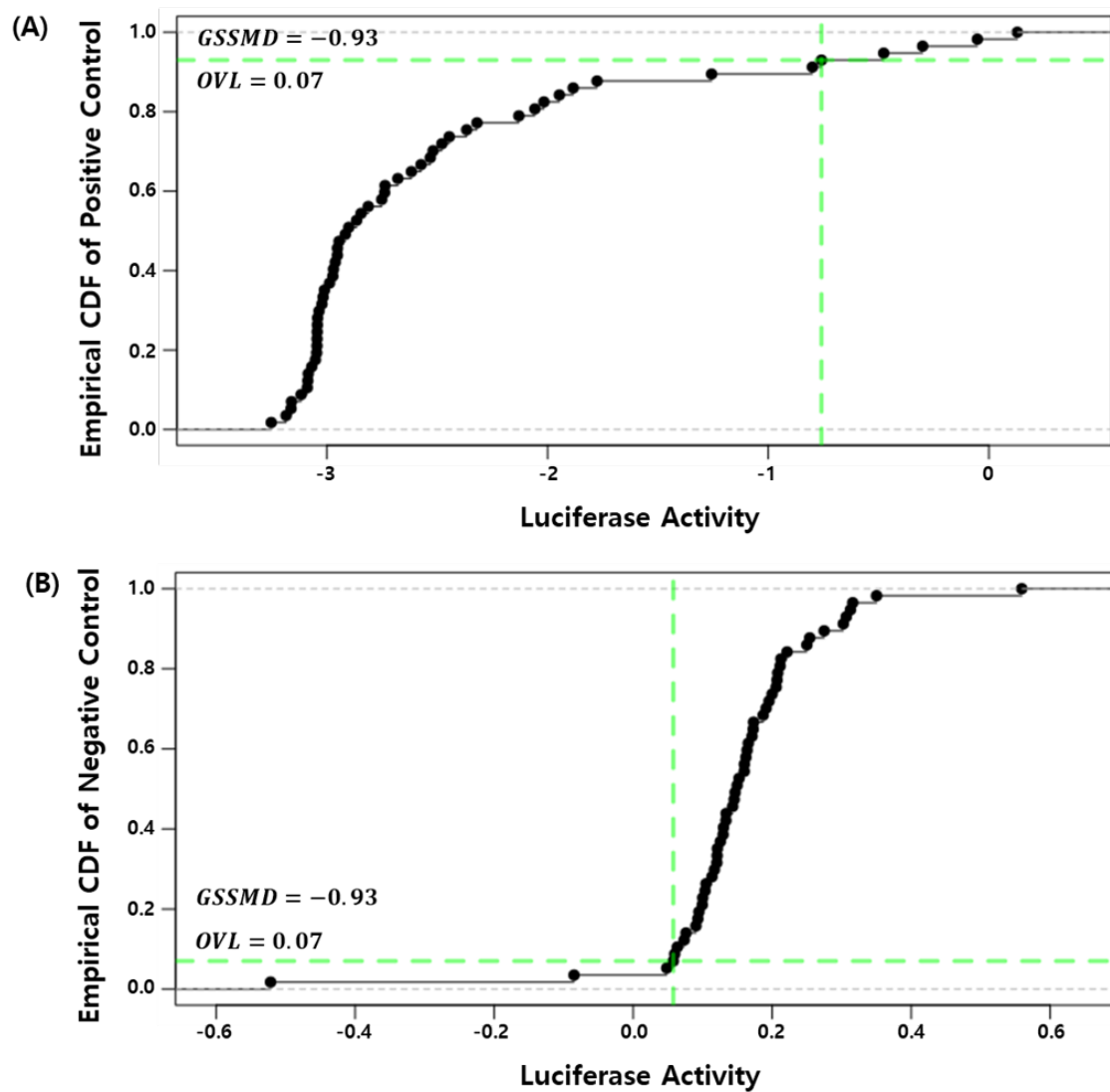


Figure S10: Selection of upper and lower bounds of threshold for hit selection. (A) Empirical CDF of positive control distribution. The horizontal green line represent 7% of ECDF corresponding to the total OVL area between positive negative distributions. (B) Empirical CDF of negative control distribution.