

Optimal Allocation of Replicates for Measurement Evaluation Studies

Stanislav O. Zakharkin^{1#}, Kyoungmi Kim², Alfred A. Bartolucci¹, Grier P. Page¹, and David B. Allison^{1,3*}

¹ *Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL 35294-0022, USA;*

² *Division of Biostatistics, School of Medicine, University of California, Davis, CA 95616-8638, USA;* ³ *Clinical Nutrition Research Center, Department of Nutrition Sciences, University of Alabama at Birmingham, Birmingham, AL 35294-3361, USA.*

Current address: Solae, LLC, St. Louis, MO 63102, USA.

Optimal experimental design is important for the efficient use of modern high-throughput technologies such as microarrays and proteomics. Multiple factors including the reliability of measurement system, which itself must be estimated from prior experimental work, could influence design decisions. In this study, we describe how the optimal number of replicate measures (technical replicates) for each biological sample (biological replicate) can be determined. Different allocations of biological and technical replicates were evaluated by minimizing the variance of the ratio of technical variance (measurement error) to the total variance (sum of sampling error and measurement error). We demonstrate that if the number of biological replicates and the number of technical replicates per biological sample are variable, while the total number of available measures is fixed, then the optimal allocation of replicates for measurement evaluation experiments requires two technical replicates for each biological replicate. Therefore, it is recommended to use two technical replicates for each biological replicate if the goal is to evaluate the reproducibility of measurements.

Key words: experimental design, measurement, microarrays, proteomics

Introduction

Optimal experimental design ensures that experiments are cost-effective, internally valid, and statistically powerful. Optimal design has particular importance for modern high-throughput technologies, such as microarrays or proteomics, because it can considerably increase the efficiency of experiments (1). In this context, it is useful to distinguish between two broad categories of experiments. We use the term “Type A” experiments to refer to the experiments conducted to address specific biological questions such as “What genes are differentially expressed or what proteins are differentially accumulated in tissue X among mice exposed versus not exposed to stimulus Y?” We use the term “Type B” experiments to refer to the experiments that are meant to evaluate aspects of the measurement technology and answer questions such as “What is the reliability of our gene expression or proteomics measurements under such-and-such

conditions?” The results of Type B studies that estimate measurement reliability are often important in helping us design optimal Type A studies.

A number of experimental layouts such as loop and common reference have been proposed, and general recommendations have been given such as necessity of proper randomization, replication, and blocking (2–4). However, while great efforts were directed toward the design for Type A studies, little attention has been given to the design for Type B experiments where the reproducibility [the terms “reliability” and “reproducibility” are often used interchangeably (5)] of measurements is assessed. Here, we address the problem of finding an optimal allocation of biological and technical replicates for such measurement evaluation studies. We provide an analytical solution based on the assumption that the total number of available measurements is fixed while biological and technical error terms are independent and normally distributed. The correctness of our analytical result is illustrated by simulations.

***Corresponding author.**

E-mail: dallison@uab.edu

Results and Discussion

Accurate information about measurement reliability (or its complement, the degree of unreliability) is often important for the design and interpretation of Type A experiments addressing specific biological questions (for example, detecting differences in gene expression levels or protein amounts between two or more conditions). Therefore, Type B experiments are needed to assess measurement reliability. While a number of experiments addressing the reproducibility of microarray and proteomics studies have been conducted (6–12), optimal design could have made such an experiment more cost-effective. For example, the stability of measurements for evaluation of two kinds of hybridization standards was analyzed using five replicates without rationalization for the number of replicates used (13). Here, we find an optimal design for Type B experiments using an analytical approach and simulations.

Briefly, we consider a scenario in which the total number of available measurements is fixed, but the number of biological replicates and the number of technical replicates are variable. Biological replicates are independent biological samples, such as rats or mice, or cell lines that have been separately subjected to biological treatments. Technical replicates are repeated measures of one biological sample. We define an allocation of replicates as optimal if it min-

imizes the variance of the ratio of technical sample variance to the total sample variance. By minimizing the variance of this ratio, we will be able to make more precise estimates of reliability defined as a ratio of a biological variance to the total variance. The results are presented in Figure 1. The variance of the ratio of technical variance to the total variance was minimal when the number of technical replicates per biological replicate equals to 2 and increased as the number of technical replicates increases. This is true for all levels of proportion of biological variance in total variance ρ^2 . Simulations were used to illustrate our analytical results. Slight differences in the shape of the plots were probably due to the fact that the approximate formulas were used in the analytical solution.

Real microarray or proteomics data are affected by multiple factors that cause variation in observed values. The choice of platform (such as cDNA arrays vs. Affymetrix), image processing algorithm (such as MAS 5.0 vs. RMA), as well as normalization and transformation techniques has a significant impact on results. Sources of variation and their magnitudes are determined experimentally and vary for different experimental systems. For example, relative contribution of biological variation ranges from ~20% to ~80% of total variation (6, 7, 10, 11). Yet, it is an important factor for experimental design, such as the planning of mRNA pooling in microarray experiments (14) and the determination of a sample size (11, 13).

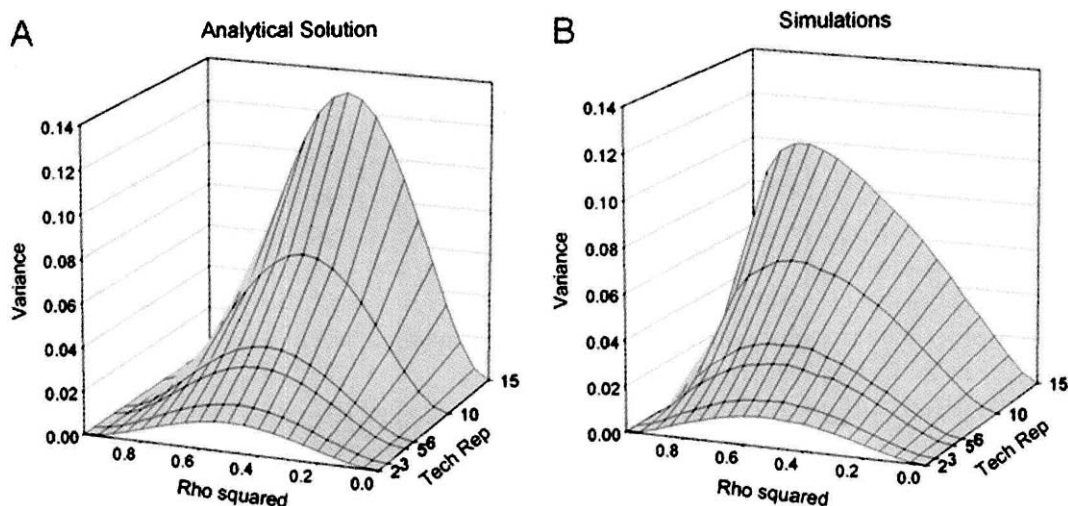


Fig. 1 The optimal number of technical replicates for measurement evaluation experiments. The graphs show the dependence of the variance of the ratio of technical variance to the total variance (Variance) on the number of technical replicates (Tech Rep) and the relative contribution of the biological variability [Rho squared (ρ^2)]. The total number of available measurements (M) is fixed in both cases as $M = 30$. **A.** Visualization of the analytical solution. **B.** Results of 10,000 times of simulations.

Our analysis uses certain assumptions to make the solution easily analytically tractable. In this study, we assume that there are two sources of variation: biological variation, which is a combination of all biological factors, and technical variation, which is a combination of all technical factors. Although this assumption is a simplification, we believe that it is reasonable and allows us to construct a realistic model.

We use a normality assumption to take advantage of approximate formulas of sample variance. While our conclusions do not depend critically on this assumption, it simplifies derivations and setup of simulations. The distribution of data remains a controversial topic in high dimensional biology. Distribution often depends on experimental system (9, 15, 16). For example, two- or three-modal distributions are not uncommon for tumor samples and may reflect underlying genetic polymorphism (17). In most cases, it is hard to reliably determine characteristics of the distribution due to small sample sizes. Hoyle *et al* (15) found in their analysis of multiple microarray datasets that distribution of spot intensities could be approximated by log-normal distribution. Giles and Kipling (16) used 59 Affymetrix GeneChips to test the normality of expression data and found that the majority of probe sets showed a high correlation with normality, with the exception of low-expressed genes. Similarly, in an experiment involving 56 independent arrays, Novak *et al* (9) found that the distribution of data processed with Affymetrix GeneChip Analysis software was very close to normal. Thus, our normality assumption is quite reasonable and can be further relaxed in future studies. In cases where the distributions of experimental data are clearly non-normal, a number of normalization and transformation methods exist to remedy the situation (18).

Decisions about the number of replicates are often influenced by available funds. Currently, in many cases, technical costs such as price of microarrays plus the cost of their processing (for example, labeling and hybridization) are significantly higher than the cost of biological samples such as mice. However, certain biological samples can be very expensive (for example, brain tumors), but it is unlikely that they will be used for evaluation of technology.

It is also noteworthy that neither the methods nor the results of our study require microarray or other high-throughput data. Our major finding—that the optimal number of technical replicates equals to two in designing studies to estimate measurement reliability—holds for any ratio of technical to biolog-

ical variance. Thus our result applies not only to the microarray context that inspired it, but also to other contexts. Additionally, our model may appear to apply only to one-channel microarrays. Yet on two-channel microarrays, our conclusions will also hold as long as we have a single value per gene per array—it could be a direct measure of intensity for single-channel arrays or a ratio (or log) difference of two measures for two-channel systems.

Based on results of this study, we give the following recommendations: use a sufficient number of samples; use appropriate normalization and transformation techniques if necessary to hold the normality assumption. It should be pointed out that our conclusion that the optimal number of technical replicates is two only applies to Type B studies to determine measurement reliability and has no direct bearing on the optimal number of technical replicates in Type A studies.

In summary, in this study we demonstrate that if the number of biological replicates and the number of technical replicates are variable, while the total number of available measurements is fixed, then the optimal allocation of replicates for measurement evaluation experiments requires two technical replicates for each biological replicate. Therefore, when investigators design measurement evaluation studies, we advocate that they use only two technical replicates for each biological replicate.

Materials and Methods

Sources of variation

The multiple factors influencing measurements in microarray and proteomics experiments could be divided into two major categories: biological variation and technical variation (1). Biological variation is caused by deviation of mRNA and protein levels within cells and among individuals (6, 7). We consider the situation where biological variation is not the major focus of studies, but rather a sampling error in traditional definition. Sample preparation, labeling, hybridization, and other factors can contribute to technical variation, which is essentially a measurement error (8–12). A common way to represent an observed value (X) is a combination of its “true” value (θ) and a measurement error (σ): $X = \theta + \sigma$ (19). We model an observed value X_i (for example, an expression level of a particular gene or an amount of a particular protein in a particular experiment after appropriate pre-

processing techniques such as normalization or transformation have been applied) as a combination of its expected value (θ), a deviation (“error term”) due to true biological variation (σ_{Bio}^2), and a deviation (“error term”) due to technical variation (σ_{Tech}^2), that is: $X_i = \theta + \sigma_{Bio}^2 + \sigma_{Tech}^2$. We assume that the biological variance is a combination of errors coming from different biological sources and the technical variance is the sum of errors from various technical sources. We define s_{Bio}^2 and s_{Tech}^2 to be estimates of σ_{Bio}^2 and σ_{Tech}^2 , respectively. We further assume that both biological and technical error terms are normally distributed: $s_{Bio} \sim N(0, \sigma_{Bio}^2)$ and $s_{Tech} \sim N(0, \sigma_{Tech}^2)$. Thus, the biological variance of a sample is s_{Bio}^2 , and the technical variance of a sample is s_{Tech}^2 . We assume that biological population variance and technical population variance are independent ($\sigma_{Bio}^2 \perp \sigma_{Tech}^2$) perhaps after suitable transformation (18), and the total population variance is the sum of biological and technical variance ($\sigma_{Tot}^2 = \sigma_{Bio}^2 + \sigma_{Tech}^2$). Assuming that there are only two sources of variation in measurement X_i , with biological and technical error terms that are independent, then the total variance in a sample is the sum of biological and technical sample variance:

$$s_{Tot}^2 = s_{Bio}^2 + s_{Tech}^2 \quad (1)$$

Reliability

Reliability is generally defined as the ratio of a “true” variation to the total variation in a set of measurements. We define the reliability, ρ^2 , as a proportion of a biological variation in the total variation, that is, as a ratio of biological population variance to the total population variance: $\rho^2 = \frac{\sigma_{Bio}^2}{\sigma_{Bio}^2 + \sigma_{Tech}^2}$. Note that our definition of reliability is essentially the intraclass correlation coefficient (18, 20). The biological population variance can be expressed as a function of the technical population variance and ρ^2 :

$$\sigma_{Bio}^2 = \frac{\rho^2}{1 - \rho^2} \sigma_{Tech}^2 \quad (2)$$

The estimate of reliability in a sample is:

$$\rho^2 = \frac{s_{Bio}^2}{s_{Bio}^2 + s_{Tech}^2} = 1 - \frac{s_{Tech}^2}{s_{Bio}^2 + s_{Tech}^2}$$

Definition of the problem

Let the total number of measurements (we define a “measurement” as the observation of a single quantitative variable on a single case) affordable in a measurement evaluation study be M , the number of biological replicates utilized be N , and the number of technical replicates per biological replicate be K , thus $M = KN$. We define a biological replicate as a biological sample, such as a rat or a tumor sample. The technical replicates are repeated measures of the same biological sample, for example, labeling and hybridization in a microarray experiment. We only consider the case of balanced design where the number of technical replicates for each biological replicate is the same. Under the assumption of the independence of biological variation and technical variation, we can estimate the biological sample variance and the technical sample variance separately. Though the ratio of technical population variance to the total population variance is a constant, the ratio of technical sample variance estimate to the total sample variance estimate, $R = \frac{s_{Tech}^2}{s_{Tot}^2}$, is a random variable whose variance depends on the sample size. The variance of the ratio R may vary as a function of the number of technical replicates. We evaluate different ways to allocate technical replicates with respect to the efficiency of an estimate of the ratio of the variance of technical replicates to the total variance. By minimizing the variance of R , we will be able to better estimate reliability in a sample, and thus better estimate reliability in a population, ρ^2 . Therefore, we define the problem of finding an optimal allocation of technical replicates for measurement evaluation experiments as follows: *For a fixed M , what value of K minimizes the variance of R ?* In the following, we provide details of analytical solution to the problem and illustrate our findings using a simulation experiment. Both approaches clearly show that the optimal value (K) of the number of technical replicates equals to 2.

Variance of the ratio R

We utilize an approximate formula for variance of the ratio of two random variables X and Y (21):

$$\text{Var} \left(\frac{X}{Y} \right) \approx \left(\frac{E(X)}{E(Y)} \right)^2 \left[\frac{\text{Var}(X)}{(E(X))^2} + \frac{\text{Var}(Y)}{(E(Y))^2} - 2 \frac{\text{Cov}(X, Y)}{E(X) E(Y)} \right] \quad (3)$$

where $E(X)$ and $E(Y)$ is the expected value of X and Y , respectively, $Var(X)$ and $Var(Y)$ is the variance of X and Y , respectively, and $Cov(X, Y)$ is the

covariance of X and Y . Equation 3 does not make assumptions about distributions of X and Y . In our case, Equation 3 can be written as:

$$Var(R) = Var\left(\frac{s_{Tech}^2}{s_{Tot}^2}\right) \approx \left(\frac{E(s_{Tech}^2)}{E(s_{Tot}^2)}\right)^2 \left[\frac{Var(s_{Tech}^2)}{(E(s_{Tech}^2))^2} + \frac{Var(s_{Tot}^2)}{(E(s_{Tot}^2))^2} - 2 \frac{Cov(s_{Tech}^2, s_{Tot}^2)}{E(s_{Tot}^2) E(s_{Tech}^2)} \right] \quad (4)$$

As we assume the independence of biological and technical sample variance, $s_{Bio}^2 \perp s_{Tech}^2$, the covariance can be expressed in variance terms:

$$Cov(s_{Tech}^2, s_{Tot}^2) = Cov(s_{Tech}^2, [s_{Bio}^2 + s_{Tech}^2]) = Cov(s_{Tech}^2, s_{Bio}^2) + Cov(s_{Tech}^2, s_{Tech}^2) = Var(s_{Tech}^2) \quad (5)$$

Equation 4 can be modified as follows (see details of derivation in the appendix):

$$Var(R) = Var\left(\frac{s_{Tech}^2}{s_{Tot}^2}\right) \approx \frac{Var(s_{Tech}^2) (E(s_{Bio}^2))^2 + Var(s_{Bio}^2) (E(s_{Tech}^2))^2}{(E(s_{Tech}^2) + E(s_{Bio}^2))^4} \quad (6)$$

We assume that both biological and technical error terms are normally distributed, $s_{Bio}^2 \sim N(0, \sigma_{Bio}^2)$ and $s_{Tech}^2 \sim N(0, \sigma_{Tech}^2)$. By the law of large numbers, the sample variance will converge to the population variance, thus the expected values of sample variance are those of population variance:

$$E(s_{Bio}^2) = \sigma_{Bio}^2 \quad (7)$$

$$E(s_{Tech}^2) = \sigma_{Tech}^2 \quad (8)$$

An approximate formula for the variance of a sample variance of a random variable X is as follows: $Var(Var(X)) \approx (\mu_4 - \mu_2^2)/S$, where μ_4 is the fourth moment and μ_2 is the second moment of X , and S is the sample size (22). In the case of a normally distributed variable $X \sim N(0, \sigma^2)$, the variance of sample variance is: $Var(Var(X)) \approx (2\sigma^4)/S$, where σ^4 is the square of the population variance of X (22). Note that the following adjusted formula gives an unbiased estimate:

$$Var(Var(X)) \approx \frac{2}{S-1} \sigma^4 \quad (9)$$

Thus, the approximate variance for estimates of the biological variance and the technical variance respectively is:

$$Var(s_{Bio}^2) \approx \frac{2}{N-1} (\sigma_{Bio}^2)^2 \quad (10)$$

$$Var(s_{Tech}^2) \approx \frac{2}{M-1} (\sigma_{Tech}^2)^2 \quad (11)$$

By substituting Equations 7, 8, 10, and 11 into Equation 6, we obtain a formula for $Var(R)$ in terms of biological and technical population variance as well as the numbers of biological and technical replicates (see details of derivation in the appendix):

$$Var(R) \approx \frac{2(\sigma_{Tech}^2)^2 (\sigma_{Bio}^2)^2}{(\sigma_{Bio}^2 + \sigma_{Tech}^2)^4} \left(\frac{1}{M-1} + \frac{1}{N-1} \right) \quad (12)$$

Finally, by substituting Equation 2 into Equation 12, we obtain a formula for the variance of the ratio of technical sample variance to the total sample variance (see details of derivation in the appendix):

$$Var(R) \approx 2\rho^4 (1 - \rho^2)^2 \left(\frac{1}{M-1} + \frac{1}{N-1} \right) \quad (13)$$

Simulations

To illustrate our analytical solution, we performed simulations using the following algorithm:

1. Sample random number B_i of K , $i = 1$ to K , from a standard normal distribution $B_i \sim N(0, 1)$. To represent "true" or biological variation, multiply each of the numbers by ρ , where ρ^2 is the ratio of biological population variance to the total population variance (see above).

2. For each of the number B_i , sample random number T_{ij} of N , $j = 1$ to N , from a standard normal distribution $T_{ij} \sim N(0, 1)$. To represent technical variation, multiply each of the numbers by $\sqrt{1 - \rho^2}$.

3. Thus, the observed value of measurements is $X_{ij} = B_i \rho + T_{ij} \sqrt{1 - \rho^2}$.

4. Compute the variance of all T_{ij} numbers (“technical sample variance”) and B_i numbers (“total biological variance”).

5. Compute the ratio $R = \frac{Var(T_{ij})}{Var(B_i) + Var(T_{ij})}$.

6. Repeat Steps 1 to 5 for desired number of times (for example, 10,000).

7. Compute the variance $Var(R)$ of the 10,000 ratios from Step 6.

Repeat Steps 1 to 7 for different N and K (given that $M = KN$ is fixed), and for different levels of ρ^2 ranging from 0 to 1.

Acknowledgements

This work was supported in part by Grants T32HL072757 and U54CA100949 from the National Institute of Health and Grant 0217651 from the National Science Foundation of USA.

Authors' contributions

SOZ performed the analysis and the simulation study. KK, AAB, and GPP contributed to discussion. DBA conceived and supervised the project. All authors contributed to the writing. The final manuscript has been read and approved by all authors.

Competing interests

The authors have declared that no competing interests exist.

References

- Churchill, G.A. 2002. Fundamentals of experimental design for cDNA microarrays. *Nat. Genet.* 32: 490-495.
- Yang, Y.H. and Speed, T. 2002. Design issues for cDNA microarray experiments. *Nat. Rev. Genet.* 3: 579-588.
- Kerr, M.K. 2003. Design considerations for efficient and effective microarray studies. *Biometrics* 59: 822-828.
- Kerr, M.K. and Churchill, G.A. 2001. Statistical design and the analysis of gene expression microarray data. *Genet. Res.* 77: 123-128.
- Quan, H. and Shih, W.J. 1996. Assessing reproducibility by the within-subject coefficient of variation with random effects models. *Biometrics* 52: 1195-1203.
- Molloy, M.P., *et al.* 2003. Overcoming technical variation and biological variation in quantitative proteomics. *Proteomics* 3: 1912-1919.
- Spruill, S., *et al.* 2002. Assessing sources of variability in microarray gene expression data. *Biotechniques* 33: 916-923.
- Brown, J.S., *et al.* 2004. Quantification of sources of variation and accuracy of sequence discrimination in a replicated microarray experiment. *Biotechniques* 36: 324-332.
- Novak, J.P., *et al.* 2002. Characterization of variability in large-scale gene expression data: implications for study design. *Genomics* 79: 104-113.
- Coombes, K.R., *et al.* 2002. Identifying and quantifying sources of variation in microarray data using high-density cDNA membrane arrays. *J. Comput. Biol.* 9: 655-669.
- Han, E.S., *et al.* 2004. Reproducibility, sources of variability, pooling, and sample size: important considerations for the design of high-density oligonucleotide array experiments. *J. Gerontol. A Biol. Sci. Med. Sci.* 59: 306-315.
- Yauk, C.L., *et al.* 2004. Comprehensive comparison of six microarray technologies. *Nucleic Acids Res.* 32: e124.
- Williams, B.A., *et al.* 2004. Genomic DNA as a cohybridization standard for mammalian microarray measurements. *Nucleic Acids Res.* 32: e81.
- Tempelman, R.J. 2005. Assessing statistical precision, power, and robustness of alternative experimental designs for two color microarray platforms based on mixed effects models. *Vet. Immunol. Immunopathol.* 105: 175-186.
- Hoyle, D.C., *et al.* 2002. Making sense of microarray data distributions. *Bioinformatics* 18: 576-584.
- Giles, P.J. and Kipling, D. 2003. Normality of oligonucleotide microarray data and implications for parametric statistical analyses. *Bioinformatics* 19: 2254-2262.
- Grant, G., *et al.* 2002. Using non-parametric methods in the context of multiple testing to determine differentially expressed genes. In *Methods of Microarray Data Analysis: Papers from CAMDA'00* (eds. Lin, S.M. and Johnson, K.F.), pp.37-55. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Shrout, P.E. and Fleiss, J.L. 1979. Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* 86: 420-428.
- Gertsbakh, I. 2003. *Measurement Theory for Engineers*, p.3. Springer-Verlag, Berlin, Germany.
- Raghavachari, M. 2004. Measures of concordance for assessing agreement in ratings and rank order data. In *Advances in Ranking and Selection, Multiple Comparisons, and Reliability* (eds. Balakrishnan, N., *et al.*), p.261. Birkhäuser Press, Boston, USA.

21. Casella, G. and Berger, R.L. 1990. *Statistical Inference*, p.331. Duxbury Press, Belmont, USA.
 22. Stuart, A. and Ord, J.K. 1994. *Kendall's Advanced*

Theory of Statistics. Vol. I. Distribution Theory (sixth edition), p.364. Hodder Arnold Publishers, London, UK.

Appendix

1. Derivation of Equation 6:

$$\begin{aligned}
 \text{Var}(R) &= \text{Var}\left(\frac{s_{Tech}^2}{s_{Tot}^2}\right) \approx \left(\frac{E(s_{Tech}^2)}{E(s_{Tot}^2)}\right)^2 \left[\frac{\text{Var}(s_{Tech}^2)}{(E(s_{Tech}^2))^2} + \frac{\text{Var}(s_{Tot}^2)}{(E(s_{Tot}^2))^2} - 2 \frac{\text{Var}(s_{Tech}^2)}{E(s_{Tot}^2) E(s_{Tech}^2)} \right] \\
 &= \left(\frac{E(s_{Tech}^2)}{E(s_{Tech}^2) + E(s_{Bio}^2)}\right)^2 \left[\frac{\text{Var}(s_{Tech}^2)}{(E(s_{Tech}^2))^2} + \frac{\text{Var}(s_{Tech}^2 + s_{Bio}^2)}{(E(s_{Tech}^2) + E(s_{Bio}^2))^2} - 2 \frac{\text{Var}(s_{Tech}^2)}{(E(s_{Tech}^2) + E(s_{Bio}^2)) E(s_{Tech}^2)} \right] \\
 &= \left(\frac{E(s_{Tech}^2)}{E(s_{Tech}^2) + E(s_{Bio}^2)}\right)^2 \left[\frac{\text{Var}(s_{Tech}^2) (E(s_{Tech}^2) + E(s_{Bio}^2))^2 + \text{Var}(s_{Tech}^2 + s_{Bio}^2) (E(s_{Tech}^2))^2}{(E(s_{Tech}^2) + E(s_{Bio}^2))^2 (E(s_{Tech}^2))^2} - \right. \\
 &\quad \left. \frac{2 \text{Var}(s_{Tech}^2) (E(s_{Tech}^2) + E(s_{Bio}^2)) E(s_{Tech}^2)}{(E(s_{Tech}^2) + E(s_{Bio}^2))^2 (E(s_{Tech}^2))^2} \right] \\
 &= \frac{1}{(E(s_{Tech}^2) + E(s_{Bio}^2))^4} \left[\text{Var}(s_{Tech}^2) \left((E(s_{Tech}^2))^2 + 2 E(s_{Tech}^2) E(s_{Bio}^2) + (E(s_{Bio}^2))^2 \right) + \right. \\
 &\quad \left. \text{Var}(s_{Tech}^2) (E(s_{Tech}^2))^2 + \text{Var}(s_{Bio}^2) (E(s_{Tech}^2))^2 - 2 \text{Var}(s_{Tech}^2) (E(s_{Tech}^2))^2 - 2 \text{Var}(s_{Tech}^2) E(s_{Bio}^2) E(s_{Tech}^2) \right] \\
 &= \frac{1}{(E(s_{Tech}^2) + E(s_{Bio}^2))^4} \left[\text{Var}(s_{Tech}^2) (E(s_{Bio}^2))^2 + \text{Var}(s_{Bio}^2) (E(s_{Tech}^2))^2 \right]
 \end{aligned}$$

2. Derivation of Equation 12:

$$\begin{aligned}
 \text{Var}(R) &\approx \frac{1}{(E(s_{Tech}^2) + E(s_{Bio}^2))^4} \left[\text{Var}(s_{Tech}^2) (E(s_{Bio}^2))^2 + \text{Var}(s_{Bio}^2) (E(s_{Tech}^2))^2 \right] \\
 &= \frac{1}{(\sigma_{Bio}^2 + \sigma_{Tech}^2)^4} \left[\frac{2}{M-1} (\sigma_{Tech}^2)^2 (\sigma_{Bio}^2)^2 + \frac{2}{N-1} (\sigma_{Bio}^2)^2 (\sigma_{Tech}^2)^2 \right] \\
 &= \frac{2(\sigma_{Tech}^2)^2 (\sigma_{Bio}^2)^2}{(\sigma_{Bio}^2 + \sigma_{Tech}^2)^4} \left(\frac{1}{M-1} + \frac{1}{N-1} \right)
 \end{aligned}$$

3. Derivation of Equation 13:

$$\begin{aligned}
 \text{Var}(R) &\approx \frac{2(\sigma_{Tech}^2)^2 (\sigma_{Bio}^2)^2}{(\sigma_{Bio}^2 + \sigma_{Tech}^2)^4} \left(\frac{1}{M-1} + \frac{1}{N-1} \right) \\
 &= \frac{2 \left(\frac{\rho^2}{1-\rho^2} \sigma_{Tech}^2 \right)^2 (\sigma_{Tech}^2)^2}{\left(\frac{\rho^2}{1-\rho^2} \sigma_{Tech}^2 + \sigma_{Tech}^2 \right)^4} \left(\frac{1}{M-1} + \frac{1}{N-1} \right) \\
 &= \frac{2(\sigma_{Tech}^2)^4 \frac{\rho^4}{(1-\rho^2)^2}}{(\sigma_{Tech}^2)^4 \left[\frac{\rho^2 + (1-\rho^2)}{1-\rho^2} \right]^4} \left(\frac{1}{M-1} + \frac{1}{N-1} \right) \\
 &= 2\rho^4 (1-\rho^2)^2 \left(\frac{1}{M-1} + \frac{1}{N-1} \right)
 \end{aligned}$$