

## STATISTICAL METHODS FOR MICROARRAY RESEARCH FOR DRUG TARGET IDENTIFICATION

David B. Allison, PhD

Department of Biostatistics, Section on Statistical Genetics, The University of Alabama at Birmingham, Birmingham, Alabama 35294-0022

**KEY WORDS:** Microarrays, empirical Bayes, mixture models, pharmacogenomics, pharmacogenetics.

Alexander Pope once wrote "Fools rush in where angels fear to tread." In more recent times, Johnny Mercer wrote "fools rush in where wise men never go. But wise men never fall in love, so how are they to know." Mercer's words are perhaps indicative of where the field of microarray design and analysis is today. There is no question that fools (including the current author) have rushed into this new and exciting field and that much of what is currently available is beneath general scientific standards for research quality and validity. However, as with any new technology, some tolerance for foolishness must be allowed so that we may come to better understand the boundaries of our existing knowledge and methodology as we allow ourselves to fall in love with exciting new technology that offers us great potential. I believe that microarray research does offer great potential, but that we must now begin increasing the statistical rigor with which we approach it in order to obtain scientifically credible insights. This paper will address some of the issues involved in this endeavor.

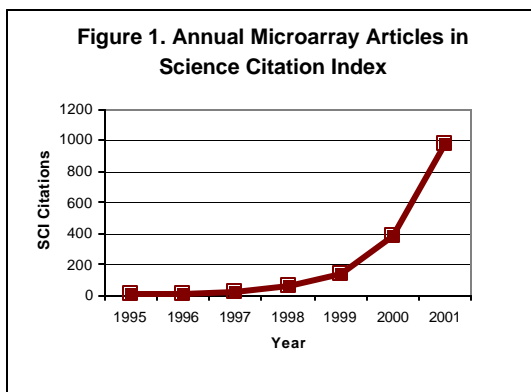
The topic specifically addressed is the use of microarray research in the field of pharmacogenomics. Pharmacogenomics can be defined as the use of genomic information to enhance the development and utilization of pharmaceuticals. The distinction between pharmacogenomics and pharmacogenetics, to the extent that there is one, is that pharmacogenetics tends to deal with one or a few genes at a time, whereas pharmacogenomics tends to involve study of many genes or the entire genome simultaneously. Although pharmacogenomics is often described as studying the genetic or genomic factors that influence response to treatment, i.e., finding genes that correlate with treatment outcome, pharmacogenomics can also be seen to include the use of genomic technology

to identify new targets, pathways, and compounds. It is to this latter purpose that I address my comments.

What are potential new drug targets? Potential new drug targets can include genes that are differentially expressed between individuals who are and are not in need of treatment for a particular disease or condition, genes that are differentially expressed when that individual is exposed to a drug known to alleviate or exacerbate the symptoms of interest, and genes that are co-expressed with other genes presumed to be involved in the systems and pathways under study. Any gene falling into any one of those categories may be a gene for which manipulation of its expression might affect disease or symptom progression. Gene expression technology may also be used in new drug candidate validation. When a new compound is identified that may be useful in treating a condition or disease, one benchmark of its potential utility might be its ability to change its overall pattern of gene expression among affected individuals to be more like that of unaffected individuals. Compounds passing this hurdle might be seen as having preliminary validation as drug candidates. This notion is predicated on the concept that gene expression patterns can be used as potential biomarkers of early or underlying response to screen for potential compounds. An ancient wisdom is captured in the phrase *in vino veritas* which means in wine there is truth. In more modern times, we might say that microarrays and related technology offer the promise of *in vitro veritas* which means in glass there is truth. That is, to the extent that we can observe *in silico* in the short term what it might otherwise take considerable time to observe *in vivo* in the long term we may be able to speed the process of drug target vetting.

An outstanding example of the thinking involved in this use of microarray technology for biomarker identification can be found in the

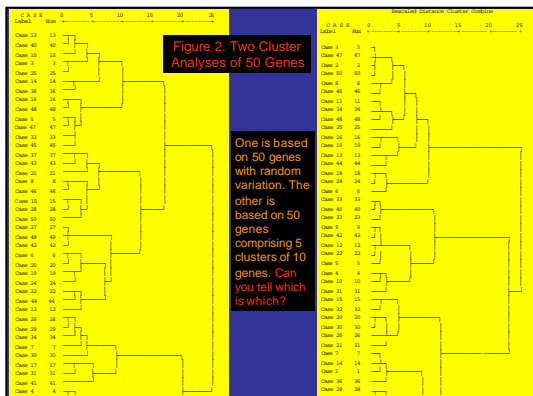
work of Lee, et al (1999). Lee, et al, studied old and young mice, as well as old mice that had been calorically restricted to find genes that were differentially expressed with age and to see which of these changes in gene expression could be "corrected" by caloric restriction, a treatment known to prolong life in multiple species including mice. At least two things are noteworthy about the paper by Lee et al. (1999). First, the findings are extremely exciting as they appear to have identified a number of genes that are differentially expressed with age, and are also affected by caloric restriction. These genes potentially allow one to study both the mechanisms of action whereby caloric restriction achieves anti-aging affects, and to identify potential targets for manipulation that might be used to slow some of the deleterious affects of aging. Also noteworthy are the strictly descriptive statistical methods used by Lee, et al. Notably absent from the paper are standard errors, confidence intervals, P-values, or any form of inferential statistic. This paper was absolutely state of the art in 1999. To a great extent, this strictly descriptive statistical approach is still in use today, but virtually all methodologists working on gene expression technology have recognized that simple descriptive statistics are insufficient and we must move beyond them to make further progress. This is critical given the enormous use to which microarray technology for gene expression research is undergoing. Figure 1 shows a plot of the number of papers from Science Citation Index, including the word microarray or microarrays from the years 1995 through 2001. As can be seen, there is a rapidly accelerating pace of research in this area that warrants the best statistical methodologies we have available. It is beyond the scope of this paper to discuss the mechanics of microarray technology. Interested readers are referred to (Cheung et al., 1999).



One of the earliest and most influential papers on the use of microarray technology was published by Eisen, et al (1998) who conducted a cluster analysis as a way of examining genome-wide expression. Specifically, the genes were clustered into different groupings using the hierarchical clustering procedure. Viewed in retrospect, the decision to use cluster analysis at that time made a good deal of sense. The authors were faced with an unusual data matrix. Specifically, if one considers that the variables under study are the expression levels of individual genes and that the cases under study are individual tumors, mice, people, etc, then the authors of that study, and indeed most microarray studies, have several orders of magnitude more variables than subjects. Conventional approaches to statistical analysis will falter in such cases because of at least two factors. First, if one conducts an analysis for each gene separately and uses a frequentist testing approach, a potential inflation of the family-wise error rate (FWER) is enormous. Second, if one tries most conventional multivariate data reduction approaches, the covariance matrix among the variables will not be positive definite, and this will create analytic difficulties. By transposing the matrix, treating the genes as cases and the cases as variables, and clustering the genes, this problem can be sidestepped. Thus, clustering the genes does not encounter the difficulties aforementioned. Subsequent to the publishing of Eisen's paper cluster analysis of microarray gene expression data became all the rage, and still is seen by many people as "the" thing to do with microarray data (Altman & Raychaudhuri, 2001). However, in this author's opinion, there are a number of issues of cluster analysis that need to be kept in mind.

Consider the two dendrograms portrayed in figure 2. One of the dendrograms is a cluster analysis of 50 genes for which expression levels were simulated such that the genes comprised five clusters, each with 10 genes per cluster. The other dendrogram depicts the results of a hierarchical cluster analysis of 50 genes where the gene expression levels were simulated from a multivariate normal distribution where the covariance matrix was an identity matrix (i.e., random noise). As can be seen, both data sets, one in which there is "true clustering" and the other in which there is not, both produce dendrograms that a creative mind can see groupings in. It is difficult to know how one

would determine which is yielding valid answers, and which is not. This is one of the great Achilles' heels of cluster analysis. Specifically, it is unclear exactly what constitutes a correct answer in cluster analysis. Moreover, and in part because of this, it is unclear when one has obtained a correct answer, how correct one's answer is, etc. I leave it as a challenge to the reader to speculate as to which one of these graphs in Figure 2 is one in which there truly is clustering, and in which one there is random noise.



These concerns about cluster analysis are not new. In 1972, Everitt (1972) stated "availability of computer calculation methods has led many psychiatrists to using the new data techniques uncritically. Spurious findings may result. The available clustering techniques should be validated as by applying them to sets of data of known structure. Most present methods may be defective." Everitt's quotation is no less true today than it was 30 years ago. Unfortunately, many scientists working in the field of microarray research are unaware of the long history of the use of cluster analysis, the problems identified, what solutions have and have not been tried, and what solutions have and have not been shown to be effective. Indeed, many writings in the field of microarray research seem to describe cluster analytic techniques as if they were a new class of techniques, not recognizing that these techniques go back at least to 1911 (Czekanoweski, 1911).

There are other concerns about cluster analysis. One is that it can be extremely computationally demanding. While this is not a criticism of the use of cluster analysis it does make the point that we must be prepared for a new level of computing support if we are to be taking on the challenge of microarray research and related genomic technologies that involve manipulating

tens of thousands of variables in methods for which we previously included far smaller data sets. In this writer's opinion, the most serious limitation of cluster analysis as a primary method in microarray research is that it does not answer the questions that are posed by the majority of applied investigators. Specifically, most applied investigators seem to have data in which microarrays have been used to measure the expression level of thousands of genes on cases from two or more groups. They wish to know which genes are differentially expressed across the two or more groups or conditions. Cluster analysis, though interesting, does not directly address this question. Elsewhere (Brand et al., 2002), we have addressed how cluster analysis may be potentially used to facilitate more traditional frequentist inference and thereby address the questions that investigators seem primarily interested in.

One can think of the major activities performed by statisticians and data analysts as involving one of five goals: measurement; design; inference; estimation; and classification. Cluster analysis can be a useful tool of classification. Through cluster analysis and related methodology, the field has spent a great deal of time working on classification. Therefore, I will not address it further herein. Our field needs to devote more attention to the other four tasks. Of these other four tasks, inference and estimation have probably garnered the most attention. The field of microarray research has spent little attention on applying formal statistical thinking to issues of measurement and design, although some work has clearly been done.

With respect to measurement, Lord Kelvin (a.k.a., William Thompson), once wrote "*in physical science, the first essential step in the direction of learning any subject is to find principles of numerical reckoning and practical methods for measuring some quality connected with it. I often say that when you can measure what you are speaking about, and express in numbers, you know something about it. But when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind; it may be the beginning of knowledge, but you are scarcely in your thoughts advanced to the state of science, whatever the matter may be.*" Lord Kelvin entreats us to consider measurement more carefully at the outset of our scientific inquiries. In this regard, a recent paper by Kuo, et al (2002)

is telling. Kuo, et al compared measurements of gene expression levels for the same genes measured on the same cases using two different microarray technologies. They conclude "corresponding measurements from the two platforms showed poor correlation. Clusters of genes and cell lines were discordant between the two technologies, suggesting that relative intratechnology relationships were not preserved." To the extent that one can see Kuo, et al's research as examining measurement reliability, the results are not encouraging. Although much work remains to be done, this suggests that our field needs more careful assessment of the measurement reliability of each of the various microarray measurement systems we utilize. Moreover, in our applied experience it appears that measurement liability varies dramatically from individual lab to individual lab. The sources of this variability are not yet fully understood and can include extraction and preparation of the tissue of interest, extraction and preparation of the mRNA from the tissue, utilization of the microarray hardware, and the specific software algorithms used to convert the measurements taken by the hardware systems to gene expression scores. Studies in the spirit of Cronbach's generalizability theory are sorely needed here (Cronbach et al., 1963).

Moving on the topic of design, I am reminded of the words of Sir Ronald Fisher who stated *"to consult the statistician after an experiment is finished is often merely to ask him to conduct a postmortem examination. He can perhaps say what the experiment died of."* As statisticians, one of the most important things that we can do to help our applied colleagues using microarray research is to encourage them to contact us for advice regarding experimental design *before* they begin the experiment. In my experience, such meetings between the applied microarray researcher and the statistician prior to initiating the experiment are more the exception than the rule, and this needs to change if we are to be in a position to provide effective support to our colleagues.

One design issue in which statisticians could have effective input concerns physical pooling of mRNA samples. Some of the microarray "chips" utilized are very expensive. This is part of the reason why some of the samples used in microarray research tends to be very small. From start to finish, the most popular type of

oligonucleotide microarray will typically cost \$1000 per subject at most universities. This does not include the cost of obtaining the mRNA sample from the subject, but only the cost of analyzing the mRNA sample once it is obtained. Given this expense, when told by a statistician that they should, for example, have several hundred subjects in each of two groups for a particular study, most applied investigators balk. A question that is sometimes asked by these investigators in return is "instead of running N microarray chips on each of the N subjects per group, can I divide the N subjects into a smaller number of groupings, physical pool the mRNA samples from all subjects within a grouping, and then analyze the pools instead?" To answer this question, let us begin with what I believe to be a critical assumption. Let us assume that the expected value of an expression measurement obtained from a pool of N subjects on a single chip equals the expected value that will be obtained from the arithmetic mean of those same N subjects measured on N separate chips. Although it is not strictly necessary for convenience, let us assume that the data (i.e., the gene expression scores) are normally distributed, that the number of subjects per condition under study is equal across conditions and evenly divisible by two within condition, and that homoscedasticity holds. Under these circumstances, one could conduct a non-pooled analysis comparing two groups or conditions. If the subjects in those two groups or conditions were independent, one could test each gene for an affect using the standard t-test (Equation 1).

$$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{N} + \frac{S_2^2}{N}}} \sim t_{2N-2} \quad (1)$$

Alternatively, one could split the N subjects from each group or condition into two pools of half-end subjects. One could then test whether the two pools from the first condition differed from the two pools from the second condition, also using a standard t-test which would take the form of Equation 2.

$$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{2} + \frac{S_2^2}{2}}} \sim t_2 \quad (2)$$

Note that in Equation 1, the variances are the variances among the individual chips, whereas in equation two the variances are the variance

among chips that contain RNA from pools from half N subjects. The non-centrality parameters for these two t-distributions are given by Equations 3 and 4.

$$d = \frac{m_1 - m_2}{\sqrt{\frac{2s_y^2}{N} + \frac{2s_e^2}{N}}} \quad (3)$$

$$d = \frac{m_1 - m_2}{\sqrt{\frac{2s_y^2}{N} + s_e^2}} \quad (4)$$

As can be seen, these non-centrality parameters are virtually identical except for the term on the right under the square root sign. In these equations,  $s_y^2$  represents the true biological variability in the dependent variable, gene expression score, and  $s_e^2$  represents the variability due to measurement error. As can be seen, Equations 3 and 4 are the same except for the fact that in equation three  $s_e^2$  will be reduced by a factor of 2/N, whereas in Equation 4 it will not be. This shows that the non-centrality parameters will be identical when there is no measurement error.

When there is a great deal of measurement error, the non-centrality parameter for the non-pooled design will be much greater and is therefore likely to offer greater power. In contrast, when measurement error is 0, barring the differences in the degrees of freedom, virtually identical power can be obtained with the pooled as with the non-pooled design at much lower cost because fewer chips will be needed. This indicates that the careful assessment of the degree of measurement error will be critical to helping us design better studies in the future, and ascertaining whether and under what circumstances pooling is an effective strategy for microarray research. This discussion also shows that, by highlighting the critical assumption that the expected value obtained from a pool of N samples of mRNA is equal to the expected value that will be obtained from the arithmetic mean and separate measurements of mRNA obtained on the same aliquots, we have highlighted a critical assumption that needs to be tested in basic bench research so that we may provide effective advice on design.

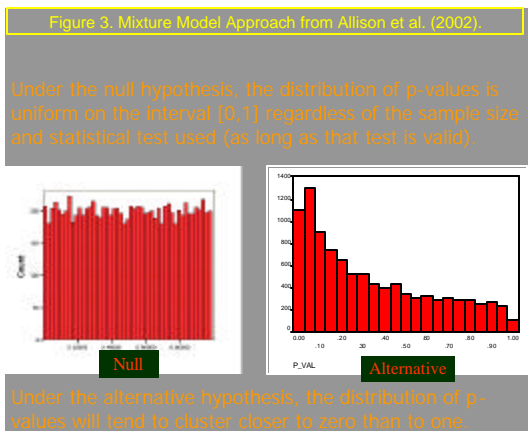
A topic that exists on the interface between design and inference concerns the use of two-

stage, or more generally multi-stage, approaches to inference when testing many hypotheses. In a recent paper, Miller et al (2001) cogently pointed out that if one tested for differences in gene expression between two or more groups or conditions for all K genes represented on an array, the plausible FWER would be enormous given that typical arrays contain many thousands of genes. Although this could easily be controlled by, for example, a Bonferroni correction, Miller et al pointed out that doing so would so dramatically reduce power that this was probably not a desirable solution. In the hopes of finding a way out of this *Scylla* and *Caribdis* problem, Miller et al suggested a two-stage approach in which one first collected data on a subset of subjects that one could afford to measure with microarrays, tested for differences between groups for all K genes on the array at some  $\alpha$  level,  $\alpha_1$ , where  $\alpha_1$  is some value greater than that which would be required by a Bonferroni correction. Some number of genes, M, will have significant effects where  $0 \leq M \leq K$ . One then gathers a second independent set of cases on which microarray measurements for gene expression are performed and test only those M genes that were found to be significant at stage one this time at  $\alpha$  level  $\alpha_2$ , where  $\alpha_2 = 0.05/M$ . Miller et al opined that this procedure would hold the Type I error rate to the nominal alpha level while making fewer Type II errors (i.e., offer greater power) than a Bonferroni correction but offered no formal evidence in support of this conjecture. Allison and Coffey (2002) formally worked out the power for these differing procedures and showed that the two-stage procedure outlined by Miller et al would not increase and, in fact, could decrease power. Although other two-stage or multi-stage procedures might be effective, the specific procedure offered by Miller et al is not advocated for use. Moreover, the paper by Miller et al points out an important caveat in the microarray field. There are many papers which offer intuitively appealing algorithms or guidelines for analyzing microarray data, but do not include rigorous analytic or simulation evidence in support of a validity of the proposed guidelines or algorithms. Readers of the microarray literature must need to learn to ask hard questions about what evidence actually supports the validity of procedures proposed for analyzing microarray data. All too often, the "evidence" supporting the validity of a proposed procedure is in fact only an illustration of the procedure with a real data set. When some

seemingly interesting findings are obtained, the authors may state that the method has been validated. However, because the true model underlying the real data set is not known, it is uncertain whether or not the correct answer was obtained. Moreover, even if one knew that the correct answer was obtained, such a demonstration provides no information about how often one would expect to obtain correct answers or how correct the average answer can be expected to be if the experiment were repeated many times. Thus readers must be cautioned that illustration of use does not indicate validity of a new procedure.

Earlier, the challenge between, on the one hand maintaining some reasonable family-wise error rate in inference and on the other hand having some reasonable power was described as a *Scylla* and *Caribdis* situation. To navigate our way through these difficult waters will require creative thinking and statistical techniques that may go beyond the typical frequentist approaches that dominate mainstream inferential statistics. One such example was provided by Allison et al (2002).

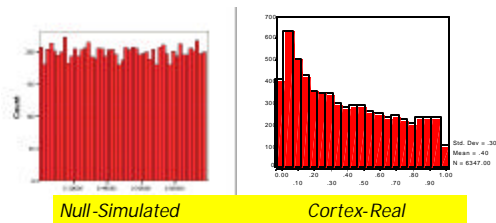
Allison et al pointed out that under the null hypothesis the distribution of P values from any study involving a large number of significant tests will be uniform on the interval [0,1] regardless of the sample size and the statistical test used as long as that test is valid. In contrast, if some of the null hypotheses are false, there will be a piling up of P values at the low end of the interval. This is illustrated in Figure 3.



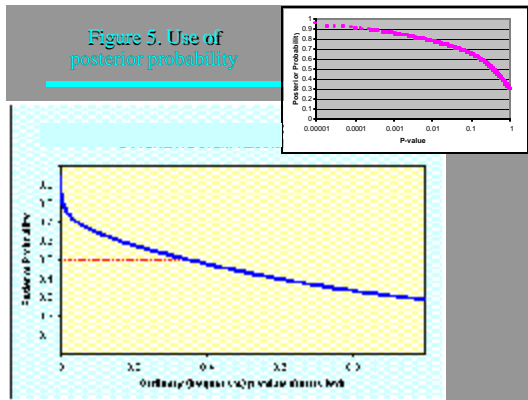
Drawing on earlier writings by other authors, Allison et al recognized that the distribution of P values under the alternative hypothesis could be modeled as a mixture of data distributions. This allows formal omnibus testing as to whether or

not there is a departure from a uniform distribution and, therefore, evidence that gene expression is differential between two or more conditions or groups. Once this is obtained, use of the fitted model coupled with Bayes rule can provide estimates of the posterior probability that a gene is a gene that is differentially expressed across the groups or conditions. Estimates can also be obtained of the total number of genes that are differentially expressed between the two or more groups or conditions. Figure 4 shows an example of this method applied to data from old calorically restricted mice as compared with old mice fed nearly ad libidum.

Figure 4. The Weindruch et al. Mouse Cortex Data (Old Ad Lib vs. Old Calorically Restricted)



The fitted mixture model indicated that roughly 29% of the genes were differentially expressed. Moreover, the mean of the beta distribution characterizing the genes for which there was differential expression was approximately 0.17. This suggests that, in this very small study of three mice/group (a typical size in microarray research) even when the null hypothesis is false, one only expects P values on the order of about 0.17. Of course this is not true for all experiments but applies to this particular data set. This illustrates that simply using 0.05 or more severe cutoffs may not be appropriate. Once the model is fit and the posterior probability obtained for each gene, these can be plotted as in figure 5. By reference to figure 5, we can help investigators decide which genes to follow up with further study. Genes with very high posterior probabilities are "safe bets." That is, they are genes for which the investigator will be unlikely to be wasting his or her time and resources if they choose to follow them up. Individual investigators can then decide what error rate they wish to tolerate, and pick genes that will give them an error rate less than or equal to that desired.



One of the questions that has received minimal attention in the field is how to determine how many subjects are needed in microarray research. Using the mixture model described above, we can apply non-parametric bootstrap techniques to extrapolate what our false positive and false negative rates would be with different sample sizes. We can then conduct power analyses that are more relevant than the experimental context of microarray research than our conventional power analyses (Gadbury et al, manuscript in preparation).

Many other methods are being vigorously investigated for their potential application to microarray research. The general concept of the false discovery rate (FDR) has been picked up by many separate investigative groups as a viable and exciting alternative to control of the FEWR. This follows from the work of Benjamini and Hochberg (1995). Several different approaches to estimating FDR's now exist. For an overview of some (but not all) FDR approaches, see Storey et al. (2002). These approaches which have the flavor of being hybrids between frequentist and Bayesian approaches seem ideally suited to microarray and other high dimensional biological research.

Finally, moving on to the topic of estimation, many investigators have also independently recognized the merit of empirical Bayes approaches for microarray research. Indeed, with microarray researchers typical studies involving estimating gene effects with thousands of genes where each estimate is based upon a relatively small amount of data, empirical Bayes technology seems ideal. Several papers have addressed this (e.g., Efron & Tibshirani, 2002) and our own research (unpublished) shows that dramatic improvements in estimation accuracy

can be obtained by very simple empirical Bayes techniques in microarray research.

In conclusion, as statisticians, we have many new and exciting frontiers ahead of us in microarray research as applied to pharmacogenomics and other areas. As a profession, we need to begin helping to increase the rigor and not just the vigor with which microarray experiments are conducted. We need to encourage our colleagues to come to us before they begin experiments, to choose numbers of experimental cases that are based upon rational and validated power and sample size estimation procedures rather than tradition or rules of thumb. We need to help educate our applied colleagues to be more skeptical consumers of the statistical literature on microarray analysis such as they can better distinguish between methods and guidelines offered that are intuitively appealing but unvalidated, and those that have a sound epistemological foundation. We need to encourage the supplementation of frequentist approaches with Bayesian approaches. Finally, we need to understand and help our colleagues to understand that the novelty of these approaches and situations makes them challenging and that appropriate time and effort needs to be provided for to allow statisticians to make an effective contribution to this exciting endeavor.

#### References

- Allison, D. B. & Coffey, C. S. (2002). Two-stage testing in microarray analysis: What is gained? *Journal of Gerontology, Biological Sciences*, 57A, B189-B192.
- Allison, D. B., Gadbury, G., Heo, M, Fernandez, J, Lee, C-K, Prolla, T. A., & Weindruch, R. (2002). A Mixture Model Approach For The Analysis of Microarray Gene Expression Data. *Computational Statistics & Data Analysis*, 39, 1-20.
- Altman RB, Raychaudhuri S. Whole-genome expression analysis: challenges beyond clustering. *Curr Opin Struct Biol*. 2001 Jun;11(3):340-7.
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple comparison problems, *J. R. Statist. Soc. B*, 57, 289-300.
- Brand, J. P. L., Page, G., & Allison, D. B. (2002). The Funnel Algorithm: Toward an

- Epistemologically Rigorous and Statistically Powerful Procedure for the Analysis of Microarray Data. Paper presented at the 2002 Annual Meeting of the Classification Society of North America June 13-16 Madison, Wisconsin, USA.
- Cheung VG, Morley M, Aguilar F, Massimi A, Kucherlapati R, Childs G. Making and reading microarrays. *Nat Genet.* 1999 Jan;21(1 Suppl):15-9.
- Cronbach, L.J., Nageswari, R., & Gleser, G.C. (1963). Theory of generalizability: A liberation of reliability theory. *The British Journal of Statistical Psychology*, 16, 137-163.
- Czekanowski, J. Objectiv kriterien in der ethnologie. *Korrespondenzblatt der Deutschen Gessellschaft fur Anthropologie, Ethnologie, und Urgeschichte*, 1911, 47, 1-5.
- Eisen, M. B., Spellman, T. P., Brown, P. O., & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *PNAS*, Vol. 95, Issue 25, 14863-14868, December 8, 1998.
- Efron B, Tibshirani R. Empirical Bayes methods and false discovery rates for microarrays. *Genet Epidemiol.* 2002 Jun;23(1):70-86.
- Everitt, B. S. Cluster analysis: A brief discussion of some of the problems. *Brit J Psychia*, 120, 1972, 143-45.
- Gadbury, G. L., Page, G. P., Edwards, J., Weindruch, R., Permana, P., Mountz, J., & Allison, D. B. Estimating Sample Size Needed to Achieve Desired True Positive, True Negative, and "Power" Rates in Microarray Research: A Parametric Bootstrap Approach. Manuscript in preparation.
- Kuo WP, Jenssen TK, Butte AJ, Ohno-Machado L, Kohane IS. Analysis of matched mRNA measurements from two different microarray technologies. *Bioinformatics.* 2002 Mar;18(3):405-12.
- Lee CK, Klopp RG, Weindruch R, Prolla TA. Gene expression profile of aging and its retardation by caloric restriction. *Science.* 1999 Aug 27;285(5432):1390-3.
- Miller RA, Galecki A, Shmookler-Reis RJ. Interpretation, design, and analysis of gene array expression experiments. *J Gerontol A Biol Sci Med Sci.* 2001 Feb;56(2):B52-7.
- Storey JD, Taylor JE, and Siegmund D. A unified estimation approach to false discovery rates. Submitted in June 2002. Technical Report #623, Dept of Statistics, University of California, Berkeley. See: <http://www.stat.berkeley.edu/users/storey/papers/623.pdf>
- Supported in part by: NIH grants R01DK56366, P30DK56336, P01AG11915, R01AG018922, P20CA093753, R01AG011653, U24DK058776 and R01ES09912, NSF grants 0217651 and 0090286 and a grant from the University of Alabama Health Services Foundation.