

Sample Size Calculation for Microarray Studies with Survival Endpoints

Sin-Ho Jung*

Department of Biostatistics and Bioinformatics, Duke University, Durham, NC, USA
Samsung Cancer Research Institute, Samsung Medical Center, Seoul, Korea

Abstract

Oftentimes, we want to discover the genes whose expression levels are associated with a time-to-event endpoint, such as progression free survival or overall survival, through microarray studies. In this case, we need to adjust the false positivity in such discovery procedure for multiplicity of the genes using a multiple testing method. The most popular multiple testing methods used for gene discovery in microarray studies are to control the false discovery rate or the family wise error rate. In this paper, we review a FDR-control method to discover the genes associated with a time-to-event outcome and propose a sample size calculation method for microarray studies designed to discover genes whose expression levels are associated the chosen time-to-event outcome. These methods can be easily modified for other types of high throughput genome projects.

Keywords: Censoring; False discovery rate; Proportional hazards model

Introduction

Microarray has been a major high-throughput assay method to display DNA or RNA abundance for a large number of genes concurrently. One of the main objectives for conducting microarray studies is to discover the genes whose expression levels are associated with a time-to-event outcome (also called a survival outcome regardless the type of event), such as time to progression or survival. Discovery of the prognostic genes should be made taking multiplicity into account, but also with enough statistical power to discover important genes successfully. Biological samples for these high-throughput experiments are costly and usually difficult to obtain, so that having too many samples will waste valuable resources, while having too few samples may lead to no meaningful scientific conclusions. Therefore, it is necessary for us to develop a sample size calculation method for gene discovery.

One important aspect in gene discovery using a microarray study is to adjust the false positivity for multiplicity of the genes. The statistical procedures with such adjustment are called multiple testing. There are two types of false positivity we control in multiple testing for gene discovery with microarray data. One is called the family wise error rate (FWER), which is defined as the probability to discover any genes when all the genes under consideration are not associated with the clinical outcome. The other is called the false discovery rate (FDR), which is defined as the proportion of unassociated genes among the discovered genes. It is well known that the multiple testing procedures controlling the FWER are so strict that they select too few genes. So, some investigators prefer the multiple testing methods controlling the FDR.

A few methods have been developed for building genomic classifiers. Most methods are focused on sample size calculations for predicting classes and are not applicable to censored outcomes data directly. Some have considered sample size estimations for gene discovery with FDR control. For example, Jung derives a sample size formula for the Storey's procedure under a weak dependence assumption [1,2]. Liu and Hwang propose similar sample size formulae that can be used for comparison of multiple independent samples. Others have also considered power and sample size calculations under FDR control [3]. Dobbin and Simon present analytical formulae for determining the number of biological replicates needed for developing a predictive classifier [4]. However,

none of these publications provides a comprehensive investigation of sample size estimation for discovery of genes whose expression levels are associated with a survival outcome.

In this paper, we consider discovering genes whose expression levels are associated with a survival outcome of patients. At first, we review multiple testing with a FDR control and Cox's regression method based on a proportional hazards model [5]. And, we propose sample size and power calculation methods for microarray studies to discover genes using Cox's regression method while controlling the FDR and for sample size calculation for gene discovery with a FWER control [6,7].

Methods

False discovery rate

Suppose that we conduct m multiple tests, of which the null hypotheses are true for m_0 tests and the alternative hypotheses are true for $m_1 (= m - m_0)$ tests. The tests declare that, of the m_0 null hypotheses, A_0 hypotheses are null (true negative) and R_0 hypotheses are alternative (false rejection, false discovery or false positive). Among the m_1 alternative hypotheses, A_1 are declared null (false negative) and R_1 are declared alternative (true rejection, true discovery or true positive). Table 1 summarizes the outcome of m hypothesis tests.

Benjamini and Hochberg defined the FDR as [8]

$$\text{FDR} = E\left(\frac{R_0}{R}\right). \quad (1)$$

Note that this expression is undefined if $\Pr(R = 0) > 0$. To avoid this issue, Benjamini and Hochberg (1995) redefined the FDR as [8]

$$\text{FDR} = \Pr(R > 0)E\left(\frac{R_0}{R} \mid R > 0\right). \quad (2)$$

These two definitions are identical if $\Pr(R = 0) = 0$, in which case we

*Corresponding author: Sin-Ho Jung, Department of Biostatistics and Bioinformatics, Duke University, Durham, NC, USA, E-mail: sinho.jung@duke.edu

Received June 06, 2013; Accepted July 19, 2013; Published July 26, 2013

Citation: Jung SH (2013) Sample Size Calculation for Microarray Studies with Survival Endpoints. J Comput Sci Syst Biol 6: 177-181. doi:10.4172/jcsb.1000114

Copyright: © 2013 Jung SH. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Accepted hypothesis			
True hypothesis	Null	Alternative	Total
Null	A ₀	A ₀	m ₀
Alternative	A ₁	R ₁	m ₁
Total	A	R	m

Table 1: Outcomes of m multiple tests.

have $FDR = E(R_0 / R | R > 0)$ (\equiv pFDR, which will be defined below).

If $m = m_0$, then $FDR=1$ by any critical value with $\Pr(R = 0) = 0$. Pointing out this issue, Storey defines the second factor in the right hand side of (2) as pFDR, [9]

$$pFDR = E\left(\frac{R_0}{R} | R > 0\right)$$

and proposes to control this quantity instead of FDR. Storey claims that $\Pr(R > 0) \approx 1$ with a large m , so that pFDR is equivalent to FDR [2]. We accept this argument in this chapter and do not distinguish between FDR and pFDR. Hence, definitions (1) and (2) are considered to be equal.

Benjamini and Hochberg (1995)[8] propose a multi-step procedure to control the FDR at a specified level. However this is known to be conservative and the conservativeness increases in m_0 [10];

Suppose that, in the j -th testing, we reject the null hypothesis H_j if the p -value p_j is smaller than or equal to $2(0; 1)$. Assuming independence of the m p -values, we have

$$R_0 = \sum_{j=1}^m I(H_j \text{ true}, H_j \text{ rejected})$$

$$= \sum_{j=1}^m \Pr(H_j \text{ true}) \Pr(H_j \text{ rejected} | H_j) + o_p(m),$$

Which equals $m_0\alpha$, where $m^{-1}o_p(m) \rightarrow 0$ in probability as $m \rightarrow \infty$ [2]. Ignoring the error term, we have

$$FDR(\alpha) = \frac{m_0\alpha}{R(\alpha)}, \tag{3}$$

Where, $R(\alpha) = \sum_{j=1}^m I(p_j \leq \alpha)$. Given α , estimation of FDR by (3) requires estimation of m_0 .

For the estimation of m_0 , Storey assumes that the histogram of m p -values is a mixture of m_0 p -values that are corresponding to the true null hypotheses and following $U(0; 1)$ distribution, and m_1 p -values that are corresponding to the alternative hypotheses and expected to be close to 0 [2]. Consequently, for a chosen constant λ away from 0, none (or few, if any) of the latter m_1 p -values will fall above λ , so that the number of p -values above λ , $\sum_{j=1}^m I(p_j > \lambda)$ can be approximated by the expected frequency among the m_0 p -values above λ from $U(0, 1)$ distribution, i.e. $m_0 / (1 - \lambda)$. Hence, given λ , m_0 is estimated by

$$\hat{m}_0(\lambda) = \frac{\sum_{j=1}^m I(p_j > \lambda)}{1 - \lambda}$$

By combining this m_0 estimator with (3), Storey (2002) obtains

$$F\hat{D}R(\alpha) = \frac{\alpha \times \hat{m}_0(\lambda)}{R(\alpha)} = \frac{\alpha \sum_{j=1}^m I(p_j > \lambda)}{(1 - \lambda) \sum_{j=1}^m I(p_j \leq \alpha)}$$

For an observed p -value p_j , Storey defines the q -value, the minimum FDR level at which we reject H_j , as [2]

$$q_j = \inf_{\alpha \geq p_j} F\hat{D}R(\alpha)$$

This formula is reduced to

$$q_j = F\hat{D}R(p_j)$$

if $FDR(\alpha)$ is strictly increasing in α , see Theorem 2 of Storey (2003) [9]. The appendix in Jung (2005) shows that this assumption holds if the powers function of the individual tests is concave in α , which is the case when the test statistics follow the standard normal distribution under the null hypotheses [1]. We reject H_j (or, equivalently, discover gene j) if q_j is smaller than or equal to the prespecified FDR level.

The independence assumption among m test statistics is loosened to independence only among m_0 test statistics corresponding to the null hypotheses [11], and to weak independence among all m test statistics [9,10]. These approaches are implemented in the statistical package called SAM [12].

Univariate Cox Regression

Let z_{ij} denote the expression level of gene $j(=1, \dots, m)$ for subject $i(=1, \dots, n)$. In order to derive the asymptotic results for sample size calculation, we assume that the expression level is bounded in Section 2. For patient i , we also observe X_i , denoting the minimum of survival time T_i and censoring time C_i together with event indicator $\delta_i = I(T_i \leq C_i)$. We assume that T_i and C_i are independent.

In order to associate the survival time with the expression level of gene j , we assume that, given z_{ij} , T_i has a hazard function of

$$\lambda_i(t) = \lambda_0(t) \exp(\beta_j z_{ij}),$$

where $\lambda_0(t)$ is an unknown baseline hazard function and β_j is the regression coefficient. By subtracting the sample mean and dividing by the standard deviation of gene j , we assume that z_{ij} has mean 0 and variance 1. In this case, β_j denotes the log-hazard ratio between two groups of patients whose expression levels of gene j are one standard deviation apart. This regression model may not hold for all genes, but still provides a reasonable measure of relationship between survival time and each gene and the test statistic presented below controls the type I error [13,14].

By Cox (1972) the partial score and information functions are given by [5]

$$U_j(\beta) = \sum_{i=1}^n \int_0^\infty \left(z_{ij} - \frac{\sum_{i'=1}^n z_{i'j} Y_{i'}(t) e^{\beta z_{i'j}}}{\sum_{i'=1}^n Y_{i'}(t) e^{\beta z_{i'j}}} \right) dN_i(t)$$

and

$$I_j(\beta) = -\frac{\partial U_j(\beta)}{\partial \beta} = \int_0^\infty \left\{ \frac{\sum_{i=1}^n z_{ij}^2 Y_i(t) e^{\beta z_{ij}}}{\sum_{i=1}^n Y_i(t) e^{\beta z_{ij}}} - \left(\frac{\sum_{i=1}^n z_{ij} Y_i(t) e^{\beta z_{ij}}}{\sum_{i=1}^n Y_i(t) e^{\beta z_{ij}}} \right)^2 \right\} dN(t)$$

respectively, where $N_i(t) = \delta_i I(X_i \leq t)$ and $Y_i(t) = I(X_i \geq t)$ denote the event and at-risk events, respectively, and $N(t) = \sum_{i=1}^n N_i(t)$. The partial maximum likelihood estimator $\hat{\beta}_j$ is obtained by solving $U_j(\beta) = 0$ usually by using the Newton-Raphson methods.

By Anderson and Gill (1982), $\hat{\beta}_j$ is a consistent estimator and

$\sqrt{n}(\hat{\beta}_j - \beta_j)$ is asymptotically normal with mean 0 and its variance can be consistently estimated by $n/I_j(\beta_j)$. The expression level of gene j is not associated with the survival time if $\beta_j = 0$, and an over-expression (under-expression) of gene j shortens the survival time if $\beta_j > 0$ (if $\beta_j < 0$). Hence, a two-sided test with marginal type I error rate α rejects $H_j: \beta_j = 0$ if

$$\left| \frac{\sqrt{n}\hat{\beta}_j}{\hat{\sigma}_j} \right| > z_{1-\alpha/2}$$

Where $\hat{\sigma}_j^2 = n/I_j(0)$ and $z_{1-\alpha}$ is the upper 100α -th percentile of $N(0,1)$. Note that $\sqrt{n}\hat{\beta}_j/\hat{\sigma}_j$ is asymptotically identical to the partial score test statistic $U_j(0)/\sqrt{I_j(0)}$.

Sample size and power calculation

Let M_0 and M_1 denote the sets of genes for which the null and alternative hypotheses are true, respectively. Note that the cardinalities of M_0 and M_1 are m_0 and m_1 , respectively. Since the estimated FDR is invariant to the order of the genes, we may rearrange the genes and set $M_1 = \{1, \dots, m_1\}$ and $M_0 = \{m_1 + 1, \dots, m\}$. For large m and under independence (or weak dependence) among the test statistics, we have [6,11].

By Storey and Tibshirani (2001) and Jung (2005), for large m and under independence (or weak dependence) among the test statistics, we have

$$\begin{aligned} R(\hat{\alpha}) &= E(R(\hat{\alpha})) + E(R(\hat{\alpha})) + o_p(m) \\ &= m_0\alpha + \sum_{j \in M_1} \xi_j(\alpha) + o_p(m), \end{aligned}$$

where $R_h(\alpha) = \sum_{j \in M_h} I(p_j \leq \alpha)$ for $h = 0, 1, \xi_j(\alpha) = P(p_j \leq \alpha)$ is the marginal power of the single α -test applied to gene $j \in M_1$. So, from (3), we have

$$FDR(\alpha) = \frac{m_0\alpha}{m_0\alpha + \sum_{j \in M_1} \xi_j(\alpha)} \quad (4)$$

by omitting the error term.

Let Z_j denote the random variable for the expression level of gene j . From Appendix, we have

$$\xi_j(\alpha) = \bar{\Phi}\left(\frac{z_{1-\alpha/2}\sigma_j - \sqrt{n}|\beta_j|}{\bar{\sigma}_j}\right),$$

where $\bar{\Phi}(\bullet) = 1 - \Phi(\bullet)$ and $\Phi(\bullet)$ denotes the cumulative distribution function of the standard normal distribution,

$$\sigma_j^{-2} = \int_0^\infty \left[\frac{E\{Z_j^2 F(t[Z_j])\}}{E\{F(t[Z_j])\}} - \left(\frac{E\{Z_j F(t[Z_j])\}}{E\{F(t[Z_j])\}} \right)^2 \right] G(t) E\{f(t|Z_j)\} dt,$$

$$\bar{\sigma}_j^{-2} = \int_0^\infty \left[E\{Z_j^2 F(t[Z_j])\} - \frac{[E\{Z_j F(t[Z_j])\}]^2}{E\{F(t[Z_j])\}} \right] G(t) dt,$$

$$F(t|z_j) = \exp\{-\lambda_0(t) \exp(\beta_j z_j)\} \text{ and}$$

$f(t|z_j) = \lambda_0(t) \exp(\beta_j z_j) F(t|z_j)$ are the survivor and probability density functions, respectively, of the survival time for a patient with gene expression level z_j and $G(t) = P(C \geq t)$. Here the expectations are taken with respect to the distribution of expression level of gene j, Z_j . These integrals are usually obtained numerically. Hence, (4) is expressed as

$$FDR(\alpha) = \frac{m_0\alpha}{m_0\alpha + \sum_{j \in M_1} \bar{\Phi}((z_{1-\alpha/2}\sigma_j - \sqrt{n}|\beta_j|)/\bar{\sigma}_j)} \quad (5)$$

Note that FDR is decreasing in n and $|\beta_j|$. Furthermore, by Jung (2005), FDR is increasing in α . If the effect sizes are equal among the prognostic genes, FDR is increasing in $\pi_0 = m_0/m$ and FDR increases from 0 to m_0/m as α increases from 0 to 1.

At the design stage of a study, m is determined when microarray chips are chosen for experiment and m_1 and $\{\beta_j, j \in M_1\}$ are projected based on experience or from pilot data if any. The only variables undecided in (5) are α and n . With all other design parameters fixed, FDR is controlled at a certain level by the chosen α level. So, we want to find the sample size n that will guarantee a certain number, say $r_1 (\leq m_1)$, of true rejections with FDR controlled at a specified level φ .

In (5), the expected number of true rejections is

$$E\{R_1(\alpha)\} = \sum_{j \in M_1} \bar{\Phi}((z_{1-\alpha/2}\sigma_j - \sqrt{n}|\beta_j|)/\bar{\sigma}_j). \quad (6)$$

In multiple testing with FDR control, $E(R_1)/m_1$ plays the role of the power of a conventional testing [15-17]. With $E(R_1)$ and the FDR level set at r_1 and φ , respectively, (5) is expressed as

$$\varphi = \frac{m_0\alpha}{m_0\alpha + r_1}.$$

By solving this equation with respect to α , we obtain

$$\alpha^* = \frac{r_1\varphi}{m_0(1-\varphi)}.$$

Given m_0, α^* is the marginal type I error level for r_1 true rejections with the FDR controlled at φ . With α and $E(R_1)$ replaced by α^* and r_1 , respectively, (6) yields an equation $h(n) = 0$, where

$$h(n) = \sum_{j \in M_1} \bar{\Phi}((z_{1-\alpha^*/2}\sigma_j - \sqrt{n}|\beta_j|)/\bar{\sigma}_j) - r_1. \quad (7)$$

We obtain the sample size by solving this equation using a numerical approach such as the bisection method. The final sample size may be chosen by rounding up the solution.

If we do not have prior information on the effect sizes, we may want to assume equal effect sizes $\beta_j = \beta_0$ for all

$j \in M_1$. Let σ^2 and $\bar{\sigma}^2$ denote σ_j^2 and $\bar{\sigma}_j^2$, respectively, obtained by replacing β_j with β_0 . Then, (7) is reduced to

$$h(n) = m_1 \bar{\Phi}((z_{1-\alpha^*/2}\sigma - \sqrt{n}|\beta_0|)/\bar{\sigma}) - r_1$$

and, by solving $h(n) = 0$, we obtain a closed form formula:

$$n = \left\lceil \frac{(\sigma z_{1-\alpha^*/2} + \bar{\sigma} z_{1-\beta^*})^2}{\beta_0^2} \right\rceil + 1, \quad (8)$$

where $\alpha^* = r_1\varphi / \{m_0(1-\varphi)\}$ and $1-\beta^* = r_1 / m_1$. Note that (8) is the conventional sample size formula when we want to detect an effect size of β_0 with power $1-\beta^*$ while controlling the type I error level at α^* .

Some practical distributional models for sample size calculation: Above sample size formula requires calculation of σ_j^2 and $\bar{\sigma}_j^2$ which depend of the distributions of survival time, censoring time and gene expression data. So, for a sample size calculation, we need to specify these distributions as accurately as possible. Suppose that there exist a database and a matching specimen bank to be used for a new microarray study. Once a sample size is calculated, we will randomly select the required number of patients from the database and specimen bank for the study. An exponential distribution can be uniquely specified by one parameter, the hazard rate, and the family of exponential distributions t real survival data relatively well. So, we often specify the survival distributions using an exponential model for which $\lambda_0(t) = \lambda_0$, a constant over time. If the expression data are standardized for a mean of 0 and a standard deviation of 1 for each gene, then we may approximate λ_0 by the hazard rate estimated from the patients in the database. If pilot data are available, we may estimate m_1 and β_j by their estimates. Otherwise, m_1 may be specified based on the experience from previous studies and β_j at a reasonable or clinically meaningful level. For example, $\beta_j = 0.5$ means that the hazard ratio of two patients whose expression levels of gene j are 0.5 standard deviation apart is $1.65 (= e^{0.5})$.

Suppose that the patients were accrued to the database with a constant rate over time period a and followed for additional period b. Then, the censoring distribution is U(b, a + b) with survivor function

$$G(t) = \begin{cases} 1 & \text{if } t \leq b \\ 1 + (b-t)/a & \text{if } b < t \leq a+b \\ 0 & \text{if } t > a+b \end{cases}$$

Another approach may be to estimate $G(t)$ from the database using the Kaplan-Meier estimator obtained by switching the role of censoring and event.

For bounded and standardized gene expression data, we may assume the truncated normal distribution with mean 0 and variance 1. The probability density function of the truncated normal distribution with mean 0 and variance 1 and truncated at $\pm\kappa$ is given as

$$f(z) = \frac{s^{-1}\phi(z/s)}{2\Phi(\kappa/s) - 1}$$

Where $\phi(\cdot)$ is the probability density function of the standard normal distribution and $s(>0)$ satisfies

$$\left[1 - \frac{2\kappa/s\phi(\kappa/s)}{2\Phi(\kappa/s) - 1}\right]s^2 = 1$$

Based on these distributional assumptions, our sample size calculation proceeds as follows:

(A) Specify the input parameters:

φ = FDR level

r_1 = number of true rejections

Distribution of standardized gene expression data

m = total number of genes for testing

m_1 = number of prognostic genes ($m_0 = m - m_1$)

$\{\beta_j, j \in M_1\}$ = regression coefficients for the prognostic genes

λ_0 = baseline hazard rate

a and b = accrual and additional follow-up periods, respectively

(B) Obtain the required sample size:

1. If the effect sizes are constant $\beta_j = \beta_0$ for $j \in M_1$, then calculate σ and $\bar{\sigma}$ and calculate n by Otherwise, calculate σ_j and $\bar{\sigma}_j$ for $j \in M_1$, and calculate n by solving

$$\sum_{j \in M_1} \bar{\Phi}((z_{1-\alpha^*/2}\sigma_j - \sqrt{n}|\beta_j|) / \bar{\sigma}_j) - r_1 = 0$$

Where $\alpha^* = r_1\varphi / \{m_0(1-\varphi)\}$.

Example 1: We assume $(m, m_1, r_1, \varphi) = (4000, 40, 1, 30, 0.1)$, $(a, b) = (5, 2)$. Suppose that the standardized expression data of each gene has a normal distribution truncated at ± 2 (i.e. $\kappa = 2$), which can be obtained by truncating $N(0, s^2)$ with $s^2 = 1.90$, and the hazard rate of a patient with gene expression z is specified as $\lambda(t|z) = \lambda_0 \exp(\beta_2 z)$ with $\lambda_0 = 1$ and $\beta_0 = 0.5$ for all $m_1 = 40$ prognostic genes. Under the assumptions on the survival and censoring distributions and the truncated expression data, we have $\sigma = 1.135$ and $\bar{\sigma} = 1.123$. Furthermore, we have

$$\alpha^* = \frac{24 \times 0.1}{3960 \times (1 - 0.1)} = 0.842 \times 10^{-3}$$

and $1 - \beta^* = 1 - 30/40 = 0.75$, so that $z_{1-\alpha^*/2} = 3.339$ and $z_{1-\beta^*} = 0.674$. Hence, from (8), the required sample size is given as

$$n = \left\lceil \frac{(1.135 \times 3.339 + 1.123 \times 0.674)^2}{0.5^2} \right\rceil + 1 = 83$$

Discussions

In this paper, we have considered discovery of genes whose expression levels are associated with a survival endpoint, and proposed sample size and power calculation methods for microarray studies that will be analyzed using the discovery method. We related the expression level of each gene with the survival outcome using a univariate Cox regression model and discovered genes by controlling the FDR to adjust for the multiplicity of genes. The proposed method is to calculate the sample size for a specified number of true discoveries while controlling the FDR at a given level. The input variables to be pre-specified are total number of genes for testing m , projected number of prognostic genes m_1 , the baseline hazard rate λ_0 , and the effectsizes β_j (the log-hazard rates between two patient groups whose gene expression levels are one standard deviation apart) for the prognostic genes. The method does not require any heavy computation, such as Monte Carlo simulations, so that we get a sample size in a second. Especially, if the effect sizes among the prognostic genes are identical, we have a closed form formula that can be calculated using a scientific calculator and a normal distribution table. The proposed method can be used to design a new study based on the parameter values estimated from pilot data.

The proposed methods can be used for other types of high-throughput biomarker study, such as genome wide association studies or gene sequencing studies, with minor modifications. The computer program for sample size calculation available from the author.

References

1. Anderson PK and Gill RD (1982) Cox's regression model for counting processes: A large sample theory. *Annals of Statistics* 10: 1100-1120.
2. Benjamini Y and Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J Royal Statistical Society Series B* 57: 289-300.
3. Cox DR. (1972) Regression Models and Life Tables (with Discussion). *J Royal Statistical Society Series B* 34: 187-220.
4. Dobbin K and Simon RM (2007) Sample size planning for developing classifiers using high-dimensional DNA microarray data. *Biostatistics* 8: 101-117.
5. Gill RD (1980) Censoring and Stochastic Integrals, *Mathematical Centre Tracts* 124. Mathematisch Centrum Amsterdam.
6. Jung SH (2005) Sample size for FDR-control in microarray data analysis. *Bioinformatics* 21: 3097-3104.
7. Jung SH, Bang H, Young S (2005) Sample size calculation for multiple testing in microarray data analysis. *Biostatistics* 6: 157-169.
8. Jung SH and Young SS (2012) Power and Sample Size Calculation for Microarray Studies. *J Biopharm Stat.* 22: 30-42.
9. Lee MLT and Whitmore GA (2002) Power and sample size for DNA microarray studies. *Stat Med* 21: 3543-3570.
10. Lin DY, Fleming TR, Gruttola VD (1997) Estimating the proportion of treatment effect explained by a surrogate marker. *Stat Med* 16: 1515-1527.
11. Storey JD (2002) A direct approach to false discovery rates. *J R Statist Soc B* 64: 479-498.
12. Storey JD (2003) The positive false discovery rate: a Bayesian interpretation and the q-value. *The Annals of Statistics* 31: 2013-2035.
13. Storey JD, Taylor JE, Siegmund D (2004) Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. *J R Statist Soc B* 66: 187-205.
14. Storey JD, Tibshirani R (2001) Estimating false discovery rates under dependence, with applications to DNA microarrays. Technical Report 2001.
15. Storey JD, Tibshirani R (2003) SAM thresholding and false discovery rates for detecting differential gene expression in DNA microarrays. In *The Analysis of Gene Expression Data: Methods and Software*, by G Parmigiani, ES Garrett, RA Irizarry and SL Zeger (editors) Springer: New York.
16. Tong T, Zhao H (2008) Practical guidelines for assessing power and false discovery rate for a fixed sample size in microarray experiments. *Stat Med* 27: 1960-1972.
17. van den Oord EJ, Sullivan PF (2003) A framework for controlling false discovery rates and minimizing the amount of genotyping in the search for disease mutations. *Hum Hered* 56: 188-199.(8).