On the solution to the Behrens-Fisher problem by the generalized p-values*

Viktor Witkovský

Abstract

In this paper we present a brief outline of the Fisher's solution to the Behrens-Fisher problem and its generalization to the problem of multiple comparisons with unequal variances by the method of generalized p-values as well as the methods to calculate the p-values required for deriving the conservative joint confidence interval estimates for the pairwise mean differences, refered to as the generalized Scheffé intervals.

Key words: Behrens-Fisher problem; Multiple comparisons; Unequal variances; Generalized p-values.
AMS classification: primary 62F04; secondary 62E15

1 Introduction

Several different solutions to the Behrens-Fisher problem have been proposed in the literature, among the others see e.g. Fisher (1935), Welch (1947), Scheffé (1970), Lee and Gurland (1975), and Robinson (1982). Based on the result of Robinson (1976), a strong theoretical support was given to the original Fisher's solution who derived the result using his theory of statistical inference called fiducial probability which is known, in other cases, to lead to paradoxes. But this is not the only way in which the solution to the problem can be justified, see e.g. Barnard (1984), Meng (1994) and Weerahandi (1995b).

In this paper, we use the equivalent solution to the Behrens-Fisher problem based on the method of generalized p-values, originally suggested by Tsui and Weerahandi (1989). This method allows to generalize the Fisher's solution to the case of multiple comparisons, see Weerahandi (1995a) and Weerahandi (1995b).

The exact p-values required for deriving the conservative joint confidence interval estimates for the pairwise mean differences, here refered to as the generalized Scheffé intervals, can be calculated by one-dimensional numerical integration.

*The research has been supported by the grant VEGA 1/7295/20 from the Scientific Grant Agency of the Slovak Republic.

2 The Fisher's distribution

The Fisher's solution is based on the $d(\theta)$ statistic,

$$d(\theta) = \frac{(\bar{Y}_1 - \bar{Y}_2) - \theta}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}, \tag{1}$$

where $\theta = \mu_1 - \mu_2$, and $\bar{Y}_i$ and S_i^2 denote the sample mean and the sample variance, respectively, of the random sample $Y_{i1}, \ldots, Y_{in_i}$, with sample size $n_i \geq 2$, from the normal population with mean μ_i and variance σ_i^2, $i = 1, 2$. We denote by $\bar{y}_i$ the observed sample mean and by s_i^2 the observed sample variance. Then, given θ, we denote by $d_{obs}(\theta)$ the observed value of the $d(\theta)$ statistic. Fisher derived, using his fiducial argument, that given s_1^2, s_2^2, and the true value of θ, the Fisher's distribution of $d(\theta)$ (or simply the distribution of d) is that of a linear combination of two independent Student's t random variables t_{f_1} and t_{f_2} with $f_1 = n_1 - 1$ and $f_2 = n_2 - 1$ degrees of freedom, i.e.

$$d \sim s_\varphi t_{f_1} - c_\varphi t_{f_2} \equiv s_\varphi t_{f_1} + c_\varphi t_{f_2}, \tag{2}$$

with

$$s_\varphi = \sqrt{\frac{(s_1^2/n_1)}{(s_1^2/n_1) + (s_2^2/n_2)}}, \quad c_\varphi = \sqrt{\frac{(s_2^2/n_2)}{(s_1^2/n_1) + (s_2^2/n_2)}}, \tag{3}$$

or equivalently, $s_\varphi = \sin(\varphi)$, $c_\varphi = \cos(\varphi)$, and $\varphi = \arctan \sqrt{(s_1^2/n_1)/(s_2^2/n_2)}$.

The Fisher's test rejects the null hypothesis $H_0 : \theta = \theta_0$ (typically $\theta_0 = 0$) against the alternative $H_1 : \theta \neq \theta_0$ if

$$|d_{obs}(\theta_0)| > \gamma_{1-\frac{\alpha}{2}}, \tag{4}$$

or if the adequate Behrens-Fisher p-value $p(d_{obs}(\theta_0))$ is sufficiently small, that is if

$$\begin{aligned} p_0 &= p(d_{obs}(\theta_0)) = 2\left[1 - \Pr\{s_\varphi t_{f_1} + c_\varphi t_{f_2} \leq |d_{obs}(\theta_0)|\}\right] \\ &= 2\left[1 - \mathcal{F}^{(d)}_{[f_1,f_2,\varphi]}(|d_{obs}(\theta_0)|)\right] < \alpha, \end{aligned} \tag{5}$$

where α is the chosen nominal significance level of the test and $\mathcal{F}^{(d)}_{[f_1,f_2,\varphi]}$ is the cdf of the random variable d. By $\gamma_{1-\frac{\alpha}{2}}$ we denote the critical value (the upper cut-off point) of the distribution of d,

$$\gamma_{1-\frac{\alpha}{2}} = \mathcal{F}^{-1}{}^{(d)}_{[f_1,f_2,\varphi]}\left(1 - \frac{\alpha}{2}\right). \tag{6}$$

Because of symmetry of the distribution we have $\gamma_{\frac{\alpha}{2}} = -\gamma_{1-\frac{\alpha}{2}}$.

The $100 \times (1-\alpha)\%$ interval estimate of θ based on $d(\theta)$ is defined as a set $\Theta_{1-\alpha} = \{\theta_0 : p(d_{obs}(\theta_0)) \geq \alpha\}$ and is given as

$$(\bar{y}_1 - \bar{y}_2) \pm \gamma_{1-\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}. \tag{7}$$

The critical values $\gamma_{1-\frac{\alpha}{2}}$ have been tabulated for some combinations of α, f_1, f_2, and φ, see e.g. Fisher and Yates (1975).

3 Solution based on the generalized p-values

The concept of generalized p-values has been introduced by Tsui and Weerahandi (1989). For the brief outline of the method see also Witkovsky (2001).

The solution to the Berens-Fisher problem by the method of generalized p-values is based on the random vector $(\bar{Y}_1, \bar{Y}_2, S_1^2, S_2^2)$ which consists a set of sufficient statistics for the parameters of the distribution. We notice, that

$$\bar{Y}_1 \sim \mathcal{N}\left(\mu_1, \frac{\sigma_1^2}{n_1}\right), \quad \text{and} \quad \bar{Y}_2 \sim \mathcal{N}\left(\mu_2, \frac{\sigma_2^2}{n_2}\right), \tag{8}$$

$$\frac{f_1}{\sigma_1^2} S_1^2 \sim \chi^2_{f_1}, \quad \text{and} \quad \frac{f_2}{\sigma_2^2} S_2^2 \sim \chi^2_{f_2}, \tag{9}$$

are mutually independent random variables.

The hypothesis of interest is

$$H_0 : \theta = \theta_0 \quad \text{against} \quad H_1 : \theta \neq \theta_0. \tag{10}$$

In this testing problem the parameter of interest is θ, $\theta = \mu_1 - \mu_2$, and (σ_1^2, σ_2^2) is the vector of nuisance parameters.

For testing H_0 and interval estimation of θ we shall define a generalized test variable $D^2(\theta_0) = D^2((\bar{Y}_1, \bar{Y}_2, S_1^2, S_2^2), (\bar{y}_1, \bar{y}_2, s_1^2, s_2^2), \theta_0, (\sigma_1^2, \sigma_2^2))$

$$D^2(\theta_0) = \frac{(\bar{Y}_1 - \bar{Y}_2 - \theta_0)^2}{(\sigma_1^2/n_1 + \sigma_2^2/n_2)} \left(\frac{\sigma_1^2}{n_1}\frac{s_1^2}{S_1^2} + \frac{\sigma_2^2}{n_2}\frac{s_2^2}{S_2^2}\right). \tag{11}$$

Notice that $D^2_{obs}(\theta_0) = ((\bar{y}_1 - \bar{y}_2) - \theta_0)^2$ does not depend on the nuisance parameters.

If $\theta = \theta_0$, the distribution of $D^2(\theta_0)$ (or simply the distribution of D^2) is given as

$$D^2 \sim \chi_1^2 \left(\frac{(s_1^2/n_1)}{\chi^2_{f_1}/f_1} + \frac{(s_2^2/n_2)}{\chi^2_{f_2}/f_2}\right), \tag{12}$$

where by χ_1^2, $\chi^2_{f_1}$ and $\chi^2_{f_2}$ we denote the independent random variables with chi-square distribution with 1, f_1 and f_2 degrees of freedom. Given $(\bar{y}_1, \bar{y}_2, s_1^2, s_2^2)$ and (σ_1^2, σ_2^2), $D^2(\theta_0)$ is stochastically increasing for $\theta > \theta_0$ and stochastically decreasing for $\theta < \theta_0$.

For given θ_0 the generalized p-value is defined as

$$p(D^2_{obs}(\theta_0)) = \Pr\left\{D^2 > D^2_{obs}(\theta_0)\right\}. \tag{13}$$

Hence, the significance test of the hypothesis H_0 is based on $p_0 = p(D^2_{obs}(\theta_0))$:

$$p_0 = 1 - \Pr\left\{\chi_1^2\left(\frac{(s_1^2/n_1)}{\chi^2_{f_1}/f_1} + \frac{(s_2^2/n_2)}{\chi^2_{f_2}/f_2}\right) \leq ((\bar{y}_1 - \bar{y}_2) - \theta_0)^2\right\}$$

$$
\begin{aligned}
&= 1 - \Pr\left\{\chi_1^2\left(\frac{f_1 s_\varphi^2}{\chi_{f_1}^2} + \frac{f_2 c_\varphi^2}{\chi_{f_2}^2}\right) \le d_{obs}^2(\theta_0)\right\} \\
&= 1 - \mathcal{F}_{[1,f_1,f_2,\varphi]}^{(F_{12}^1)}(d_{obs}^2(\theta_0)), \qquad (14)
\end{aligned}
$$

where $\mathcal{F}_{[1,f_1,f_2,\varphi]}^{(F_{12}^1)}$ is the cdf of the random variable

$$
F_{12}^1 \sim \chi_1^2\left(\frac{f_1 s_\varphi^2}{\chi_{f_1}^2} + \frac{f_2 c_\varphi^2}{\chi_{f_2}^2}\right). \qquad (15)
$$

We reject H_0 if the p-value p_0 is small (smaller than the nominal significance level α), or if

$$
d_{obs}^2(\theta_0) > \mathcal{F}^{-1}{}_{[1,f_1,f_2,\varphi]}^{(F_{12}^1)}(1-\alpha). \qquad (16)
$$

Note, that (14) is equivalent to

$$
p_0 = 1 - \Pr\left\{-\frac{d_{obs}^2(\theta_0)}{\chi_1^2} + \frac{f_1 s_\varphi^2}{\chi_{f_1}^2} + \frac{f_2 c_\varphi^2}{\chi_{f_2}^2} \le 0\right\}, \qquad (17)
$$

i.e. the p-value p_0 is a function of the cdf of a linear combination of independent inverted chi-square random variables. Exact values can be calculated by the method suggested by Witkovský (2001), which requires one-dimensional numerical integration.

We note, that the p-value (14) is equal to the Behrens-Fisher p-value (5). If we denote

$$
\gamma_{12,1-\alpha}^1 = \sqrt{\mathcal{F}^{-1}{}_{[1,f_1,f_2,\varphi]}^{(F_{12}^1)}(1-\alpha)}, \qquad (18)
$$

then $\gamma_{12,1-\alpha}^1 = \gamma_{1-\frac{\alpha}{2}}$, where $\gamma_{1-\frac{\alpha}{2}}$ is defined by (6), and the $100\times(1-\alpha)\%$ interval estimate of θ based on the generalized p-values is given as

$$
(\bar{y}_1 - \bar{y}_2) \pm \gamma_{12,1-\alpha}^1\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}. \qquad (19)
$$

4 Generalization to the multiple comparisons

Consider now $k \ge 2$ independent samples from normal populations, $Y_{i1}, \ldots, Y_{in_i}$, with sample sizes n_i and with possibly unequal means μ_i and variances σ_i^2, $i = 1, \ldots, k$.

Let $\bar{Y}_i = (1/n_i)\sum_{j=1}^{n_i} Y_{ij}$ be the sample means and let $S_i^2 = (1/f_i)\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_i)^2$ be the sample variances, $f_i = n_i - 1$. Then $\bar{Y}_i \sim \mathcal{N}(\mu_i, \sigma_i^2/n_i)$ and $f_i S_i^2/\sigma_i^2 \sim \chi_{f_i}^2$ are mutually independent random variables. Further, let $\bar{y}_i$ be the observed value of $\bar{Y}_i$ and let s_i^2 be the observed value of S_i^2, $i = 1, \ldots, k$. Hence, the observed value of the random vector $\bar{Y} = (\bar{Y}_1, \ldots, \bar{Y}_k)'$ is $\bar{y} = (\bar{y}_1, \ldots, \bar{y}_k)'$, and the observed value of the random vector $S^2 = (S_1^2, \ldots, S_k^2)'$ is $s^2 = (s_1^2, \ldots, s_k^2)'$.

Let $\mu = (\mu_1, \ldots, \mu_k)'$, $\sigma^2 = (\sigma_1^2, \ldots, \sigma_k^2)'$, and let

$$R = \begin{pmatrix} 1 & -1 & 0 & \ldots & 0 \\ 1 & 0 & -1 & \ldots & 0 \\ \vdots & & & & \vdots \\ 1 & 0 & 0 & \ldots & -1 \end{pmatrix}, \tag{20}$$

R being an $(\kappa \times k)$-matrix, $\kappa = k-1$. We will denote $\theta = R\mu$, i.e. $\theta = (\theta_1, \ldots, \theta_\kappa)'$ with $\theta_i = \mu_1 - \mu_{i+1}$. Then, additionally assuming that $R\bar{y} - \theta_0 \neq 0$, the significance test of the null hypothesis on all contrasts $H_0 : \lambda'\theta = \lambda'\theta_0$ for all $\lambda \in \mathcal{R}^\kappa$ against the alternative $H_1 : \lambda'\theta \neq \lambda'\theta_0$ for some $\lambda \in \mathcal{R}^\kappa$, with $\theta_0 = (\theta_{01}, \ldots, \theta_{0\kappa})'$, (typically $\theta_0 = 0$), will be based on the generalized test variable $F^\kappa(\theta_0) = F^\kappa(\bar{Y}, S^2, \bar{y}, s^2, \theta_0, \sigma^2)$, where

$$F^\kappa(\theta_0) = \frac{(R\bar{Y} - \theta_0)'[RVR']^{-1}(R\bar{Y} - \theta_0)}{(R\bar{y} - \theta_0)'[RWR']^{-1}(R\bar{y} - \theta_0)}, \tag{21}$$

with

$$V = \operatorname{diag}\left(\frac{\sigma_1^2}{n_1}, \ldots, \frac{\sigma_k^2}{n_k}\right) \text{ and } W = \operatorname{diag}\left(\frac{\sigma_1^2}{n_1}\frac{s_1^2}{S_1^2}, \ldots, \frac{\sigma_k^2}{n_k}\frac{s_k^2}{S_k^2}\right). \tag{22}$$

Note that, $F^\kappa_{obs}(\theta_0) = F^\kappa(\bar{y}, s^2, \bar{y}, s^2, \theta_0, \sigma^2) = 1$. Under H_0 we have $(R\bar{Y} - \theta_0) \sim \mathcal{N}(0, RVR)$, and hence, $(R\bar{Y} - \theta_0)'[RVR']^{-1}(R\bar{Y} - \theta_0) \sim \chi^2_\kappa$. On the other hand, the denominator of $F^\kappa(\theta_0)$ is stochastically independent on the nominator and $(R\bar{y} - \theta_0)'[RWR']^{-1}(R\bar{y} - \theta_0) \sim q(\bar{y}, s^2, \theta_0, \chi^2_{f_1}, \ldots, \chi^2_{f_k})$ with

$$RWR' \sim \begin{pmatrix} \frac{(s_1^2/n_1)}{\chi^2_{f_1}/f_1} + \frac{(s_2^2/n_2)}{\chi^2_{f_2}/f_2} & \frac{(s_1^2/n_1)}{\chi^2_{f_1}/f_1} & \cdots & \frac{(s_1^2/n_1)}{\chi^2_{f_1}/f_1} \\ \frac{(s_1^2/n_1)}{\chi^2_{f_1}/f_1} & \frac{(s_1^2/n_1)}{\chi^2_{f_1}/f_1} + \frac{(s_3^2/n_3)}{\chi^2_{f_3}/f_3} & \cdots & \frac{(s_1^2/n_1)}{\chi^2_{f_1}/f_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{(s_1^2/n_1)}{\chi^2_{f_1}/f_1} & \frac{(s_1^2/n_1)}{\chi^2_{f_1}/f_1} & \cdots & \frac{(s_1^2/n_1)}{\chi^2_{f_1}/f_1} + \frac{(s_k^2/n_k)}{\chi^2_{f_k}/f_k} \end{pmatrix}. \tag{23}$$

If $\theta = \theta_0$, the distribution of $F^\kappa(\theta_0)$ (or simply the distribution of F^κ) is given as

$$F^\kappa \sim \frac{\chi^2_\kappa}{q(\bar{y}, s^2, \theta_0, \chi^2_{f_1}, \ldots, \chi^2_{f_k})}, \tag{24}$$

and does not depend on the vector of nuisance parameters σ^2.

For given θ_0 the generalized p-value is defined as

$$p(F^\kappa_{obs}(\theta_0)) = \Pr\left\{F^\kappa > F^\kappa_{obs}(\theta_0)\right\}. \tag{25}$$

Hence, the significance test of the hypothesis H_0 is based on $p_0 = p(F^\kappa_{obs}(\theta_0))$:

$$p_0 = 1 - \Pr\left\{\frac{\chi^2_\kappa}{q(\bar{y}, s^2, \theta_0, \chi^2_{f_1}, \ldots, \chi^2_{f_k})} \leq 1\right\}, \tag{26}$$

and the $100 \times (1-\alpha)\%$ confidence set for θ based on the generalized p-values is defined as a set

$$\Theta^{\kappa}_{1-\alpha} = \{\theta_0 : p(F^{\kappa}_{obs}(\theta_0)) \geq \alpha\}. \tag{27}$$

From (26) we have

$$\begin{aligned} p_0 &= 1 - \Pr\left\{\chi^2_\kappa \leq q(\bar{y}, s^2, \theta_0, \chi^2_{f_1}, \dots, \chi^2_{f_k})\right\} \\ &= 1 - \Pr\left\{\chi^2_\kappa \leq (R\bar{y} - \theta_0)'[RWR']^{-1}(R\bar{y} - \theta_0)\right\} \\ &\leq 1 - \Pr\left\{\chi^2_\kappa \leq \frac{(\lambda'(R\bar{y} - \theta_0))^2}{\lambda' RWR'\lambda}\right\} = p_0^\lambda, \end{aligned} \tag{28}$$

for each fixed non-zero vector $\lambda = (\lambda_1, \dots, \lambda_\kappa)'$, $\lambda \in \mathcal{R}^\kappa$. In general,

$$p_0 \leq p_0^{\lambda_*} = \inf_{\lambda \in \mathcal{R}^\kappa} p_0^\lambda, \tag{29}$$

note that the equality $p_0 = p_0^{\lambda_*}$ holds true if $\kappa = 1$. The p-value p_0 can be calculated numerically by a method which requires in general κ-dimensional numerical integration, see Weerahandi (1995b). However, for any $\lambda \in \mathcal{R}^\kappa$ we can calculate

$$p_0^\lambda = 1 - \Pr\left\{-\frac{\left(\sum_{i=1}^{\kappa} \lambda_i((\bar{y}_1 - \bar{y}_{1+i}) - \theta_{0i})\right)^2}{\chi^2_\kappa} + \sum_{i=0}^{\kappa} \lambda_i^2 \frac{(s^2_{1+i}/n_{1+i})}{\chi^2_{f_{1+i}}/f_{1+i}} \leq 0\right\}, \tag{30}$$

where $\lambda_0^2 = \left(\sum_{i=1}^{\kappa} \lambda_i\right)^2$. Note, that the p-value p_0^λ can be represented as a function of the cdf of a linear combination of independent inverted chi-square random variables and the exact values can be calculated by the method suggested by Witkovský (2001).

Consider now the problem of pairwise multiple comparisons. Then for $i = 1, \dots, k$ and $j = i+1, \dots, k$ we define $p_0^{(ij)} = p_0^{\lambda^{(ij)}}$, i.e.

$$\begin{aligned} p_0^{(ij)} &= 1 - \Pr\left\{\chi^2_\kappa \left(\frac{(s_i^2/n_i)}{\chi^2_{f_i}/f_i} + \frac{(s_j^2/n_j)}{\chi^2_{f_j}/f_j}\right) \leq \left((\bar{y}_i - \bar{y}_j) - \theta_0^{(ij)}\right)^2\right\} \\ &= 1 - \Pr\left\{\chi^2_\kappa \left(\frac{f_i s^2_{\varphi_{ij}}}{\chi^2_{f_i}} + \frac{f_j c^2_{\varphi_{ij}}}{\chi^2_{f_j}}\right) \leq d^2_{ij,obs}\left(\theta_0^{(ij)}\right)\right\} \\ &= 1 - \mathcal{F}^{(F^\kappa_{ij})}_{[\kappa, f_i, f_j, \varphi_{ij}]}\left(d^2_{ij,obs}\left(\theta_0^{(ij)}\right)\right), \end{aligned} \tag{31}$$

where $\theta_0^{(ij)} = \mu_{i0} - \mu_{j0}$, $\mathcal{F}^{(F^\kappa_{ij})}_{[\kappa, f_i, f_j, \varphi_{ij}]}$ is the cdf of the random variable

$$F^\kappa_{ij} \sim \chi^2_\kappa \left(\frac{f_i s^2_{\varphi_{ij}}}{\chi^2_{f_i}} + \frac{f_j c^2_{\varphi_{ij}}}{\chi^2_{f_j}}\right), \tag{32}$$

and further,

$$d_{ij,obs}\left(\theta_0^{(ij)}\right) = \frac{(\bar{y}_i - \bar{y}_j) - \theta_0^{(ij)}}{\sqrt{s_i^2/n_i + s_j^2/n_j}}, \tag{33}$$

Diagnosis	n_i	$\bar{y}_i$	s_i^2
CONTR	11	6.7194	3.7092
AML	25	11.6970	40.4761
BALL	8	8.2017	7.8417
CLL	7	5.4709	0.9178
CML	7	16.5356	33.8495
NHL	11	15.2078	64.0564
TALL	12	10.7675	21.3017

Table 1: The mean basal DNA damage of patients with hematological malignacies

and $s^2_{\varphi_{ij}} = \sin^2(\varphi_{ij})$, $c^2_{\varphi_{ij}} = \cos^2(\varphi_{ij})$, and $\varphi_{ij} = \arctan\sqrt{(s_i^2/n_i)/(s_j^2/n_j)}$. Note that the p-value $p_0^{(ij)}$ given by (31) is a generalization of the Behrens-Fisher p-value p_0 given by (5) and (14).

The conservative joint $100 \times (1-\alpha)\%$ confidence interval estimates for the pairwise mean differences $\theta^{(ij)} = \mu_i - \mu_j$, refered to as the generalized Scheffé intervals, are given as

$$(\bar{y}_i - \bar{y}_j) \pm \gamma^{\kappa}_{ij,1-\alpha}\sqrt{\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}}, \tag{34}$$

$i = 1, \ldots, k$, and $j = i+1, \ldots, k$.

The critical values

$$\gamma^{\kappa}_{ij,1-\alpha} = \sqrt{\mathcal{F}^{-1}{}^{(F^{\kappa}_{ij})}_{[\kappa, f_i, f_j, \varphi_{ij}]}(1-\alpha)}, \tag{35}$$

can be calculated numerically by solving the equation

$$1 - \alpha = \Pr\left\{-\frac{\left(\gamma^{\kappa}_{ij,1-\alpha}\right)^2}{\chi^2_{\kappa}} + \frac{f_i s^2_{\varphi_{ij}}}{\chi^2_{f_i}} + \frac{f_j c^2_{\varphi_{ij}}}{\chi^2_{f_j}} \leq 0\right\}. \tag{36}$$

5 Example

Bačová *et al.* (1998) used single cell gel electrophoresis to evaluate the level of DNA damage measured in per cent of tail DNA in peripheral blood, bone marrow, and lymphatic node cells of patients with acute lymphoblastic leukemia ALL, (ALL of T-cell subtype is denoted by TALL, ALL of early B-cell subtype is denoted by BALL), acute myeloid leukemia AML, chronic lymphocytic leukemia CLL, chronic myeloid leukemia CML and non-Hodgkin's lymphoma NHL. The level of DNA damage is to be compared with the level of basal DNA damage in control group CONTR, represented by healthy donors.

The mean basal DNA damage increased in order

$$\text{CLL} < \text{BALL} < \text{TALL} < \text{AML} < \text{NHL} < \text{CML},$$

Contrast	$\bar{y}_i - \bar{y}_j$	Individual comparisons *p*-value	95% CI		Multiple comparisons *p*-value	95% CI	
CONTR-AML	-4.9776	**0.0016	-7.8955	-2.0596	*0.0935	-10.4225	0.4674
CONTR-BALL	-1.4822	0.2433	-4.1410	1.1766	0.9476	-6.7265	3.7621
CONTR-CLL	1.2485	0.1074	-0.3122	2.8093	0.7990	-1.7766	4.2737
CONTR-CML	-9.8162	**0.0045	-15.3350	-4.2973	*0.0922	-21.1516	1.5193
CONTR-NHL	-8.4884	**0.0062	-14.0082	-2.9685	0.1663	-19.3035	2.3267
CONTR-TALL	-4.0481	**0.0166	-7.2406	-0.8555	0.3474	-10.1992	2.1031
AML-BALL	3.4954	*0.0501	-0.0025	6.9932	0.6371	-3.1253	10.1160
AML-CLL	6.2261	**0.0001	3.4599	8.9923	**0.0103	1.0441	11.4081
AML-CML	-4.8386	*0.0982	-10.7875	1.1103	0.7503	-16.7696	7.0923
AML-NHL	-3.5108	0.2277	-9.4697	2.4480	0.9461	-14.9971	7.9754
AML-TALL	0.9295	0.6292	-2.9869	4.8459	0.9997	-6.4290	8.2880
BALL-CLL	2.7307	**0.0355	0.2368	5.2247	0.4598	-2.2671	7.7286
BALL-CML	-8.3340	**0.0116	-14.1723	-2.4957	0.2110	-20.1458	3.4779
BALL-NHL	-7.0062	**0.0226	-12.8482	-1.1642	0.4000	-18.3539	4.3416
BALL-TALL	-2.5659	0.1630	-6.2975	1.1658	0.8936	-9.7417	4.6100
CLL-CML	11.0647	**0.0024	-16.5103	-5.6191	*0.0535	-22.3012	0.1718
CLL-NHL	-9.7369	**0.0025	-15.1810	-4.2929	*0.0829	-20.4394	0.9656
CLL-TALL	-5.2966	**0.0026	-8.3531	-2.2401	*0.0946	-11.2340	0.6408
CML-NHL	1.3278	0.7088	-6.2335	8.8890	0.9999	-13.4038	16.0594
CML-TALL	5.7681	*0.0605	-0.3187	11.8549	0.6143	-6.3927	17.9289
NHL-TALL	4.4403	0.1407	-1.6539	10.5346	0.8651	-7.2980	16.1787

Table 2: Analysis of the DNA damage data. By $*$ we denoted the *p*-values less than 0.1, by $**$ we denoted the *p*-values less than 0.05.

what correlated with survival prediction of patients with particular hematological disease. A large heterogeneity was found in the level of basal DNA damage among patients with AML and NHL.

Heparinized blood, bone marrow samples and lymphatic nodes of 70 patients with hematological malignacies and of 11 healthy donors, respectively, were obtained[1] from the National Cancer Institute and from the Department of Pediatric Oncology of University Children's Hospital, Bratislava, Slovak Republic. For illustration purposes, we present simplified analysis. For each patient we calculated the mean level of DNA damage. Based on such data we calculated the sample means and the sample variances for each type of diagnosis, including control group of healthy donors. The data in agregated form are presented in the Table 1. In the Table 2 we present the results of individual as well as multiple comparisons. The individual comparisons lead to the problem of comparing equality of the mean basal DNA damage for each two diagnosis separately, in this case we set $\kappa = 1$. The multiple comparisons give us an answer to the question if the method based on measuring basal DNA damage could be a potential diagnostic tool for the patients with hematological maligancies, in this case we set $\kappa = 6$. The answers to

[1]The data were kindly provided by Gabriela Bačová of the Cancer Research Institute, Slovak Academy of Sciences, Bratislava, Slovak Republic.

the above questions are readily available from the results presented in the Table 2. However, in this paper, we leave this discussion without any further comment.

References

[1] Bačová, G., Gábelová, A., Babušíková, O., and D. Slameňová (1998). The single cell gel electrophoresis: A potential tool for DNA analysis of the patients with hematological maligancies. *Neoplasma*, 45, 349–359.

[2] Barnard, G.A. (1984). Comparing the means of two independent samples. *Applied Statistics*, 33, 266–271.

[3] Fisher, R.A. (1935). The fiducial argument in statistical inference. *Annals of Eugenics*, 6, 391–398.

[4] Fisher, R.A. and F. Yates (1975). *Statistical Tables.* Longman, London.

[5] Lee, A.F.S. and J. Gurland (1975). Size and power of tests of equality of two normal populations with unequal variances. *Journal of the American Statistical Association*, 70, 933–941.

[6] Meng, X.-L. (1994). Posterior predictive p-values. *Annals of Statistics*, 22, 1142-1160.

[7] Robinson, G.K. (1976). Properties of Student t and of the Behrens-Fisher solution to the two means problem. *Annals of Statistics*, 4, 963–971.

[8] Robinson, G.K. (1982). Behrens-Fisher problem. *Encyclopedia of Statistical Sciences*, 9 volumes plus Supplement, Wiley, New York, 205–209.

[9] Scheffé, H. (1970). Practical solutions of the Behrens-Fisher problem. *Journal of the American Statistical Association*, 65, 1501–1508.

[10] Tsui, K.W. and S. Weerahandi (1989). Generalized p values in significance testing of hypotheses in the presence of nuisance parameters. *Journal of the American Statistical Association*, 84, 602–607.

[11] Weerahandi, S. (1995a). ANOVA under unequal error variances. *Biometrics*, 51, 589–599.

[12] Weerahandi, S. (1995b). *Exact Statistical Methods for Data Analysis.* Springer-Verlag, New York.

[13] Welch, B.L. (1947). The generalization of 'Student's' problem when several different population variances are involved. *Biometrika*, 34, 28–35.

[14] Witkovský, V. (2001). Exact tests of variance components using generalized p-values. In: (I. Horová Ed.), *Folia, Fac. Sci. Nat. Universitatis Masarykianae Brunensis, Mathematica 9, Proceedings of the Summer School Datastat 99*, 119–125.