Letter to the Editor

The anti-atheroslerotic effect of *Allium sativum*: Statistics re-evaluated

The authors responsible for the statistics have reevaluated the study 'The antiatherosclerotic effect of *Allium sativum'* published in Atherosclerosis [144 (1999) 237–249] with different statistical methods other than applied in the original presentation. Regretfully, we have to inform you that in part deviating findings resulted which we would like to specify here.

The statistical problem in the calculation of P-values in Table 8 (original publication) lies in the neglect of the dependence between the four measurements of plaque volumes over time per probationer. Therefore, the columns of P- and SEM-values in Table 8 are not correct. Based on fixed follow-up times, the application of the Mann–Whitney U-test (with respect to the quotient of measurement 4/measurement 1) as originally intended in its one-tailed form in the study design, would have been a simple solution of this problem. The difference between the final examination (measurement 4) and the baseline examination (measurement 1) is calculated and compared between the groups. This has also the advantage that the difference could correct individual influences from the probationer himself. Because only two-tailed tests are recommended since 5th February 1998 ([1], Section 5.5), two-sided tests throughout are collected in Table 1.

Table 1
Synopsis of P-values calculated by different methods for the simple comparison of measurement 4 versus 1 differences between placebo and verum groups as well as for significant modifications of the regression lines under garlic treatment (cf. original Figs. 6–8 and Table 8, time-dependent evaluation)

Method	Total collective	Men	Women
Mann–Whitney U-test			
Difference	0.0641	0.4102	0.0124
Quotient	0.0424	0.3952	0.0107
Linear mixed effects model			
Per protocol	0.1427	0.7271	0.0391
Intention to treat	0.3122	0.3443	0.0218
F-test			
Combination	0.1127	0.9296	0.0140

On the other hand, the authors wanted to compare the time course of the plaque volume in the placebo and verum group over the four examinations in the follow-up. In view of the interdependence of the four investigations in each participant as well as the inhomogeneity in axis intercept and slope of the individual regression lines, so-called 'linear mixed effects models' are recommended in the statistics to treat these problems.

Retrospectively, within the frame of such models the statistical procedure is a modified split-plot analysis of variance, where the blocks are regarded as being chosen at random from a population of blocks and the plots within the blocks give the repeated measurements within the primary sampled unit [2]. The existence of a modification of the regression line under garlic treatment is depicted in Table 1.

The linear mixed effects model is a generalization of the classical covariance analysis. Usually, in covariance analysis the influence of a continuous covariable is adjusted through a regression model. On the other hand, one can diminish the error in estimating the regression coefficients by estimation of a specific regression line for each level of the factor 'treatment' (adjusted regression [3,4]). Through separation of variance components, this leads to F-test statistics. F-test results were also included in Table 1.

The P-values of the age-dependent evaluation procedure in Table 8 did not reach significance by the methodological approaches utilized here.

We used the same linear mixed effects model for 'intention to treat' analysis to study whether a model with treatment effect explains the observed data better than a model without treatment effect. The likelihood ratio test gave the results as indicated in Table 1. Again, for females a significant effect is shown.

The difference in age between placebo and verum group is more marked in women than in men (Mann–Whitney U-test, women: $P = 0.0073$, men: $P = 0.0678$). The missing correlation, however, between age and change in plaque volume over 4 years in women (Spearman rank correlation; difference: $\rho = -0.20$, $P = 0.1707$, quotient: $\rho = -0.18$, $P = 0.2158$) can be used as an argument that a distortion through age is not likely in this subgroup, and thus a significant result

Table 2
Slopes (m) and axis intercepts (b) of the regression lines $y = mx + b$ for placebo derived from a linear mixed effects model[a]

Gender	Parameter	Per protocol	Intention to treat
Men + women	m (mm³/month)	0.12 ± 0.08	0.11 ± 0.07
	b (mm³)	33.12 ± 2.27	32.17 ± 1.72
	Δ (mm³/month)	-0.16 ± 0.11	-0.11 ± 0.10
Men	m (mm³/month)	0.04 ± 0.09	0.04 ± 0.08
	b (mm³)	37.55 ± 2.81	35.91 ± 2.10
	Δ (mm³/month)	-0.04 ± 0.12	0.03 ± 0.13
Women	m (mm³/month)	0.30 ± 0.14	0.30 ± 0.14
	b (mm³)	23.23 ± 3.41	23.24 ± 2.67
	Δ (mm³/month)	-0.37 ± 0.17	-0.38 ± 0.18

[a] Values are given with standard error. Per protocol probationers: 105 men, 47 women; per protocol observations: 420 men, 188 women; intention to treat: 198 men, 82 women; intention to treat observations: 546 men, 237 women. The change in slope under treatment is indicated by Δ. 'Per protocol' values are compared with 'intention to treat' values.

between placebo and verum group can be ascribed to the garlic treatment. An average increase of 0.30 mm³ per month in plaque volume is estimated (Table 2). This is a population value. For the single individual a value resulted that is normally distributed, with a mean value of 0.30 and a standard deviation of 0.51. This increase is undone by garlic treatment. The reduction of plaque volume is estimated to be 0.37 mm³ per month (Table 2) or 74.9% after 4 years ($PV_0 = 23.7$ mm³, cf. original publication). Furthermore, under garlic treatment an effective regression of plaque volume by 0.07 mm³/month, or 14.2% related to the observation time of 4 years, has been found.

Comparing the slopes m of regression lines calculated according to the linear mixed effects model (Table 2) with those originally published (Table 8), almost identical values ensued. The reduction in plaque volume over a 4-year period was estimated to be 23.1% in the total collective (-0.16 mm³/month) and 5.1% in men (-0.04 mm³/month), respectively. While this effect in men + women varied between significant, borderline and trend, the reduction in plaque volume was non-significant in men, significant, however, in women (cf. Table 1). Having noticed practically no changes of plaque volume in men after the age of sixty, the conclusion is obvious that the observation period of four

years had been too short to be able to detect a significant garlic effect and/or that it would have been more appropriate to include male probationers between 20 and 60 years of age into the study.

Conclusions

Utilizing different statistical methods as well as considering the per protocol and the 'intention to treat' analysis, one arrives in principle at similar conclusions. Significant effects of garlic were confirmed in women and there was a clear trend in the total collective of men + women, the effect in men, however, turned out to be non-significant.

Acknowledgements

References

[1] ICH Expert Working Group. ICH harmonised tripartite guideline. Statistical principles for clinical trials. Stat Med 1999;18:1905–42.
[2] Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics 1982;38:963–74.
[3] Wernecke K-D. Angewandte Statistik für die Praxis. Bonn, Paris, Reading, Menlo Park: Addison-Wesley, 1995.
[4] Wernecke K-D, Ulrich WD. Application of covariance analysis to ophthalmologic research. Biomed J 1978;20:107–14.

8 February 2000

G. Siegel, D. Klüßendorf
Institute of Physiology,
Biophysical Research Group,
University Clinic Benjamin Franklin
of the Free University of Berlin,
Arnimallee 22, D-14195 Berlin,
Germany
E-mail: siegel@mail.grumed.fu-berlin.de