Multi-stage designs for the measurement of the preventive effect of sunscreens against immune suppression due to sunlight

Franz Quehenberger
Institute for Medical Informatics, Statistics and Documentation
franz.quehenberger@kfunigraz.ac.at

Peter Wolf
Department of Dermatology

Karl Franzens University, Graz, Austria

Handout at the poster session of the RSS annual meeting
Reading, September 2000

1 Introduction

Animal experiments indicate that the protective effect of chemical UVfilters against immune suppression due to sunlight might be quite different from their protective effect against sunburn. Unfortunately deficiencies of the immune system have been shown to be related to the development of skin malignancies. The immune system's ability to acquire contact allergies is suppressed by exposure to sunlight. In this study the protective effect of 3 different sunscreen preparations against this suppression is investigated.
Originally the study was designed for a pilot stage and 2 study stages. Due to high error variance and heteroscedasticiy the final model was not available before the last volunteer was in the study. Therefore formal optimal experimental design methods could only be applied retrospectively. This analysis will be presented here.

2 The study set-up

First the *minimum erythema dose* (**MED**) was assessed for each volunteer. Then a sunscreen was applied on the left buttock in a standardized way. This should protect against the subsequent application of a multiple (0 to 3) of the MED multiplied by the sun protection factor. This quantity will also be called MED below. *Sensitisation* was performed 24 hours later by application of a very effective but rarely used allergenic substance (DNCB) (Fig. 1). The other preparations had a sun protection factor of about 5.

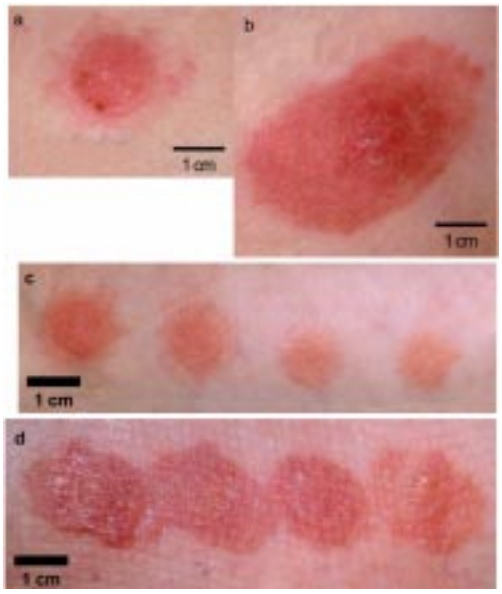

Figure 1: Slight (a) and intense (b) erythema caused by sensitisation through DNCB. 3 weeks later the allergic reaction can be tested on the left upper arm. The skin reaction for 4 different doses of DNCB after 49 (c) and 72 hours (d).

The *allergic* (i.e. immune) *reaction* was tested 3 weeks later the with DNCB on the upper arm (challenge). The skin reaction was assessed by the *dermal oedema* (change in skin thickness) after 49 and 72 hours, as measured by an ultrasound device.
The *treatment groups* (A, B, C) were defined by 3 kinds of chemical UVfilters applied before UV exposure. Sunlight doses were multiplied with the sun protection factor of the UVfilter. Group A was the control group and received sun protection factor 1.
Each volunteer could undergo sensitisation only once. Due to restrictions only 80 persons could be enrolled. After the suspicion arose that women could show extra variability due the menstrual cycle, they (n = 7) were excluded from the analysis and recruitment was restricted to males after the pilot stage.
After inspection of the results of the pilot stage (20 volunteers, control treatment) with sunlight doses from 0 to 3 MED, the dose range was restricted to 0 to 1.5 for the next 30 volunteers (10 per treatment group). Although initially the use of optimal experimental designs was intended the next 30 doses were arranged again relying on heuristical insight gained from [3], as at this stage it was not possible to decide on which statistical model to use (Fig. 2).

3 The model

The objective of the study is to determine the MED dose at which the immune reaction (oedema size) is decreased by half (**ED50**).
If the protective effect against immune suppression is similar to the protection

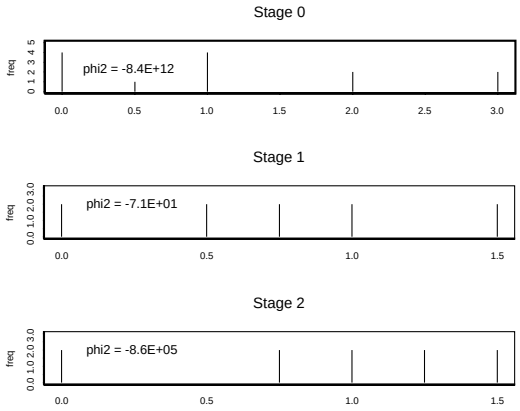

Figure 2: MED doses by stage. Stage 0 was the pilot with treatment A only. The priors for ϕ_2 came from the final model (see text).

against sunburn one would expect no differences between treatment groups. Moreover one is interested in the position of ED50 relative to MED 1.
In order to answer these questions a *4-parameter logistic model* with transformation of both dependent (with the power -0.5) and the independent (log transform) data was fitted for the dependence of the skin thickness on the logarithm of the MED doses:

$$T(y) = T\left\{\text{logistic}\left(\frac{\log(X)-\log(a)}{b}\right) * c + d\right\} + \epsilon$$

T	a normalizing transformation ($u^{-1/2}$)
y	skin thickness
logistic(x)	$1/\{1+\exp(x)\}$
X	sunlight dose in MED units
ϵ	independent errors with expectation 0
a	ED50, the dose that causes 50% reduction of skin reaction (one parameter for each treatment group).
b	slope parameter
c	change in skin thickness, oedema
d	normal skin thickness

The normalizing transformation was chosen by fitting a regression line to log transformed absolute residuals and log(MED) and inspection of the corresponding scatter plot. However, variances were not stabilized completely. Outlying observations were caused for example by blisters.
As parameter d was almost equal to the mean skin thickness before the challenge (1.14 mm), the oedema sizes (differences) were fitted in the sequel.
High variability of oedema sizes and scarce design points for low doses caused divergence of ED50, slope parameters and / or oedema versus infinity in analysis by stage 1.

Table 1: Nonparameteric bootstrap estimate of the final model ($n = 72, R = 200$).

Parameter	estimate	median bias	std. error
ED50A	0.62	-0.04	0.21
ED50B	0.67	-0.00	0.24
ED50C	0.83	0.01	0.30
b	-0.35	0.01	0.35
c	1.14	0.01	0.23

The final model fit is shown in Fig. 3 and Table 1. Bootstrap samples showed marked correlation (max $= 0.81$) of model parameters. This correlation can be reduced (max $= 0.52$) by restricting the parameter space of the slope parameter b to $[-0.4, 0.2]$.

4 Multistage designs

In optimal experimental design usually some criterion on the Fisher information matrix of the experiment is maximized. In nonlinear problems this criterion in general depends on the parameters under study. In multistage designs the design of further stages is based on parameter estimates that were obtained in previous stages.
The necessity to know the parameter values can be overcome by Bayesian design methods [1]. The posterior distribution of the previous stages is he prior distribution of the next stage. The optimal Bayesian design η maximizes the following quadratic loss criterion:

$$\phi_2(\eta) = -\int \operatorname{tr}\left\{ A(\theta)\,[nI(\theta,\eta)]^{-1} \right\} p(\theta)\,\mathrm{d}\theta.$$

Above $I(\theta,\eta)$ is the expected Fisher information matrix, depending on the model parameters θ and the design measure η. p is the prior information on the parameters. Weighting of parameters is achieved by the matrix A. If e.g. the information on only one parameter is to be maximized one chooses $A = cc^T$, whereby c is an indicator vector.
Optimum Bayesian designs can be calculated easily with software by M. Clyde [2] within XLISPSTAT [4].

5 Post hoc analysis of the study design

The univariate nonparametric bootstrap distributions from the final model were used as prior distributions for finding the optimal Bayesian designs (more detailed: rescaled χ^2distributions with the same first and second moments, discretized at 4 points). The standard deviation of the slope parameter b had to be set to 0.1, as the optimizer failed to converge with the standard error of the estimate obtained from the study.

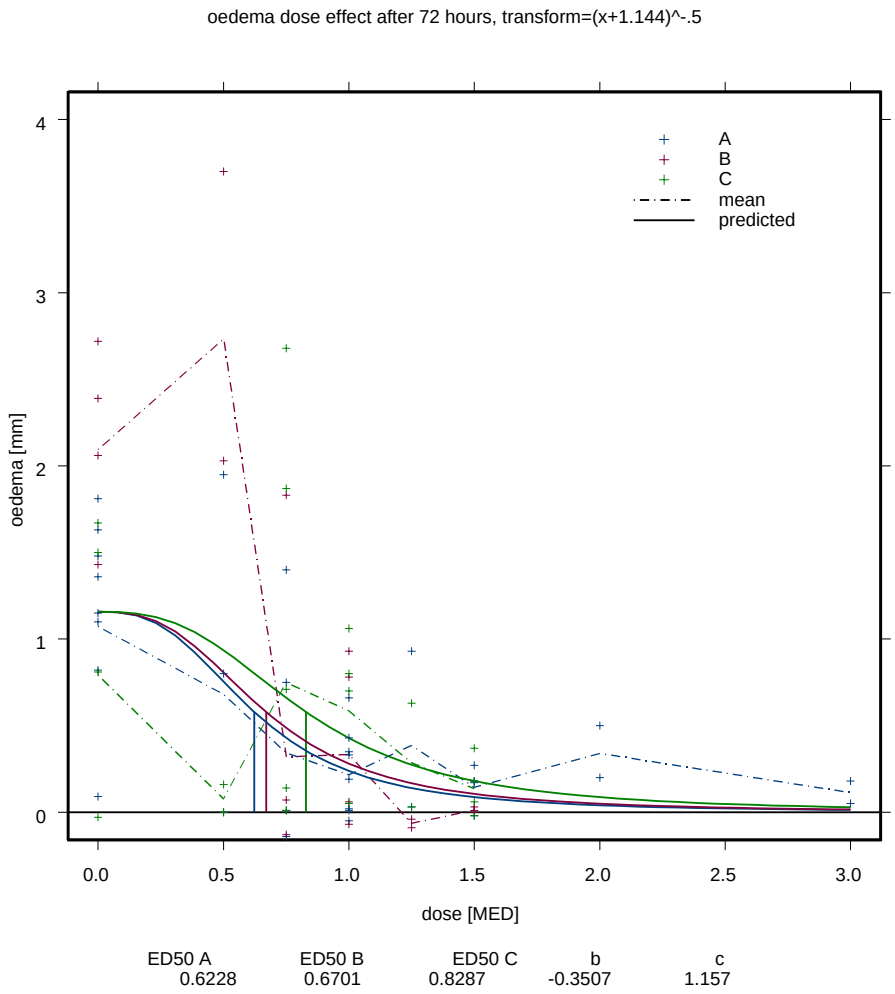

Figure 3: Figure 3: Measurements of oedema due to DNCB challenge of three types of sunscreens ($n = 72$). Means on transformed scale (dotted lines) and model estimates (solid lines).

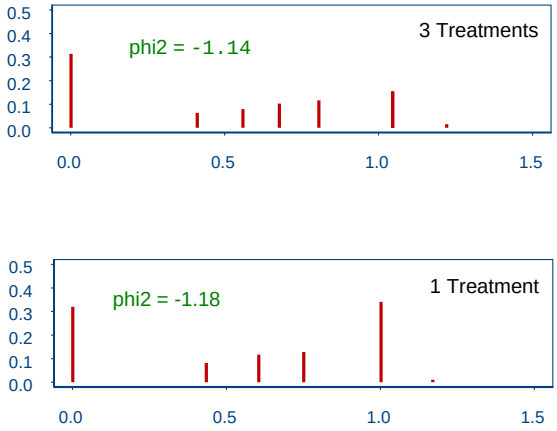

Figure 4: Optimal Bayesian design based on the final model.

Dummy coding of the treatment groups would have required 3 dimensions for the design space. As the contribution of the location parameter a of each group to the Fisher information is identical, the design was computed assuming one group and design space dimension one. Results for 3 treatment groups were obtained by multiplying the first derivative of the mean function with respect to parameters b and c by $\sqrt{3}$.
The post hoc ϕ_2-criterion of the 3 study stages is given in Fig. 2. The design choice for stage 2 was motivated by the optimal designs presented in [3] as no reliable model was available at that time. However, assuming $a = 1.1$ instead of $a = 0.71$ (an estimate in from these stage 1 models) and $SE_a = 0.4$ instead of $SE_a = 0.5$ one gets $\phi_2 = -40$. So the choice of design was not so far from optimal. Unfortunately no convergence of the optimizing design algorithm could be achieved for realistically flat priors.
The optimal post hoc designs assuming 1 and 3 treatment groups are presented in Fig. 5. In order to make comparisons ϕ_2 was calculated for the model with 3 treatment groups in both designs. So there is almost no difference in efficiency between the optimal designs (3.3%). However, the ϕ_2-criterion of the corresponding study with 1 treatment group 31.19 times the ϕ_2-criterion of a study with 3 treatment groups. So the efficiency gain by combining 3 treatments in one study is about 0.19.

6 Conclusions

As is well known optimal design has to rely on a fixed model. This prevented the application of optimal multistage designs in our experiment. A final model could be given in the last stage. Insights gained from the design literature were valuable. There is still inconvenience with the model because of high correlations of some parameter estimates.
Post hoc analysis of the final model revealed substantial dependence of the de-

sign efficiency on the prior distribution of the parameters. Thorough choice of priors will be advisable for the Bayesian design of future experiments. This property of the model also promises high efficiency gains by multistage designs.
The number of volunteers per substance can be reduced considerable by testing more than one substance simultaneously.
The program provided by M. A. Clyde is a valuable tool for computing optimal Bayesian designs. Due to the clear software design modifications were made with ease.

References

[1] K Chaloner and I Verdinelli. Bayesian experimental design: A review. *Statistical Science*, 10:273–304, 1995.

[2] M A Clyde. An object-oriented system for Bayesian nonlinear design using XLISP-STAT. Technical Report 587, University of Minnesota, 1993.

[3] S Lauer, J Timmer, D van Calker, D Maier, and J Honerkamp. Optimal weighted bayesian design applied to dose-response-curve analysis. *Communications in Statistics, Part A-Theories and Methods*, 26:2879–2903, 1997.

[4] L Tierney. *Lisp-Stat: an object-oriented environment for statistical computing and dynamic graphics.* Wiley, New York, 1990.