Assessing maximally allowed concentration; application to the regulation of PNOS

G. SALANTI, K. ULM

Institut für Medizinische Statistik und Epidemiologie der TU München

Klinikum rechts der Isar

Ismaninger Str. 22

81675 Muenchen

Deutschland

Introduction

A threshold value, in Germany called MAK value (maximum workplace concentration) is defined as the maximum concentration of a chemical substance not have known adverse effects on the health of the employee nor cause unreasonable annoyance even when the person is repeatedly exposed during long periods, given a 40-hour working week. As a rule, the MAK value refers to the average concentration obtained by integrating the concentrations determined during a period of up to one working day or shift (DFG, 2002).

In assessing a threshold value (a MAK value) the most sensitive endpoint has to be taken into account. Regarding the exposure to dust, which can cause unspecific effects on the respiratory organs, chronic bronchitic reaction (= CBR) was used. The decision whether there was such a reaction or not was based on a classification tree combining an anamnestic questionnaire and lung function tests. In order to assess a threshold value in this situation and to adjust for the effect of other important factors, like age, a statistical model is necessary.

Statistical methods

General considerations

To analyse the impact of a continuous variable on a defined response a model has to be used, in which the form of the relationship is specified. The general model using the variables “dust concentration” c and “time since first exposure” t was as follows:

$$\text{logit } p = \alpha + f_1(c) + f_2(t) \tag{1}$$

The data of the CBR-study were divided into several subsamples with respect to plant and smoking. Several models were applied to analyze the DFG-Chronic Bronchitis study. The models differ with respect to the choice of the functions $f_1(c)$ and $f_2(t)$.

Logistic regression and threshold value estimation

The simplest model is the *linear* logistic regression that assumes linear relationships. That means both functions in formula (1) are linear (e. g. $f_1(c) = ß_c \cdot c \quad f_2(t) = ß_t \cdot t$).

This function can be extended for estimating a threshold value τ

$$f_1(c) = \begin{cases} 0 & c \le \tau \\ \beta_c (c - \tau) & c > \tau \end{cases} \tag{2}$$

In case of a threshold value there is the restriction that the regression line has to be constant up to the break point. Basically the idea is to introduce an additional parameter into the regression equation, the threshold value. A more formal description of the statistical method is given by Ulm (1991).

One problem is related to the outcome. The relationship can result in a local decrease in the risk with increasing exposure. Such a result could be seen in some subsamples in the DFG Bronchitis Study where the risk of the disease is lower than the baseline risk over some range of the lower dust concentration.

This phenomenon was also observed in other studies. There are several explanations to observe a decrease in the risk. First the workers with zero or very low exposure might be different from the exposed workers, or there is a selection bias. Second this result may be explained by chance.

Isotonic regression

The isotonic transformation and its features

Alternative to the logistic model a nonparametric approach, isotonic regression, has been used (Robertson, Wright & Dykstra, 1988). Isotonic regression has some advantages compared to parametric models. No specific assumptions for the form of the dose-response relationship apart from monotonicity are required.

Isotonic regression is fitted by the Polled Adjacent Violators Algorithm (PAVA), an algorithm widely used and easy to apply. Isotonic framework has been well established by Robertson et al. and gained further attention since then.

Isotonic regression also provides a test for trend which is powerful, stable and independent of any monotonic transformation of the predictors (Chuang-Stein et al., 1997), a feature that is not provided by commonly used test (as the Cochran- Armitage test) or the logistic regression.

Technical details of the PAVA and isotonic test for trend are presented in the appendix.

Isotonic regression and threshold estimation

Isotonic regression proceeds by splitting the predictors in constant risk groups and yields a small set of cutpoints. These cutpoints are candidate threshold locations (Morton-Jones et al., 2000). Note that isotonic regression is a tool for modeling and testing, that can be applied in different forms: as univariate regression, as surface-estimator or to extend generalized additive models.

Regarding applications already presented for assessing a threshold value for dust, isotonic regression has been fitted in two dimensions applying an isotonic-surfaces model. In past papers (Ulm 1999) the following model has been used:

$$p = f(c,t) \qquad (3)$$

with $f(c_1,t_1) \le f(c_2,t_2)$ for $c_1 \le c_2$ or $t_1 \le t_2$.

For estimating $f(c,t)$ the data had to be grouped in order to reduce the dimensionality. In lack of an appropriate test procedure for assessing threshold values, a risk increase of 5%

compared to baseline was used, which is of course somewhat arbitrary. In contrast to that, in recent papers (Ulm 1999 and Salanti & Ulm 2001) isotonic regression was extended and new procedures were introduced in order to assess a threshold value.

Here we will present how an *additive isotonic model* can be used to assess threshold values. It has been first introduced by Bacchetti (1989) and improved by Morton-Jones (2000). The model takes the usual additive form described earlier (formula (1)).

$$\text{logit}\, p = \alpha + f_1(c) + f_2(t) \qquad (1)$$

where now $f_i(c)$ and $f_2(t)$ are isotonic transformation functions (step functions). Here, the data need no longer to be grouped into certain categories.

The effect of dust concentration can be tested by comparing the values of the likelihood functions of model (1) with and without the dust in the model ($f_1(c) = 0$). There is so far no distributions theory available for the corresponding likelihood ratio test, so conditional permutations as described by Salanti & Ulm (2001) need to be used to assess significance (see appendix). Note that this test is a multivariate test for trend for dust (adjusted for variable time since first exposure), and therefore provides an important tool in establishing monotonic dose-response relationship.

As result of the estimation procedure, $f_1(c)$ amalgamates certain levels of dust concentration in order to give one risk estimate for each category. Note that this estimation as well as the permutation test are adjusted for other predictors (e. g. time) included in model.

Once the model is established and a dose-response relationship is proven, - i.e. the permutations test is significant, a threshold can be assessed. The procedure can be described as follows:

1. The categories defined for dust from the partial fit in model (1) i.e. $f_1(c)$ are lumped together starting from the two lowest categories

2. The change in the goodness of fit (= deviance) is estimated

3. If the loss in the fit is significant (one-sided χ^2-distributed) a threshold is assessed. Otherwise the procedure continues (go to step 1) until a significant change in the deviance is reached.

Results

Data from 5578 workers of three different plants (Moers, Munich and Saarbrücken) were available. A detailed description of the data can be found in the monograph of the study (DFG-report 1978) or in Ulm et al. (1996). The three plants had a mixture of dust, mainly from iron, steel, foundry and engineering.

Logistic regression and isotonic-surfaces models

Within this paper we restrict ourselves to the assessment of the threshold for inhalable dust. The results of this approach consisted the basis for assessing the MAK-values and are presented in table 1. The estimation of the threshold regarding isotonic regression has been performed using a 5% excess in the risk. The threshold values are varying between 3.8 and 20.6 mg/m^3 (Ulm et al., 1996).

Table 1: Results used for assessing the threshold value for total dust

Non-smokers	sample		
type of regression	Moers	Munich	Saarbrücken
logistic	20.6	8.0	7.5
Isotonic-surfaces	2.5	6.0	–

Smokers	sample		
type of regression	Moers	Munich	Saarbrücken
logistic	18.0	3.8	13.8
Isotonic-surfaces	4.5	5.0	3.5

Which of these estimates should be taken as MAK-values? The depends also on the policy of the MAK-commission. Given an estimated interval within which the threshold lies, the lowest value can be used in order to fulfil the definition of the preamble (s. Greim at al., 2002). However, in the final report, the MAK commission decided to set the threshold at 4 mg/cm³.

Isotonic regression

The improved version of isotonic regression was used to estimate a threshold value. In three of the subsamples isotonic regression leads to a significant effect for dust concentration (s. Table 2). In the estimation of a threshold value for dust we proceed as described in the previous section. The estimated threshold values are between 2.09 and 11.09 mg/m³. In three of the subsamples a significant effect of dust was to be seen. In two of the three subsamples with a significant dust effect (the smokers of Munich and Moers) the old and the new threshold values are close.

Table 2: Results from the additive isotonic model for assessing threshold values (new results)

Samples	p-value for the effect of dust	threshold value for total dust
Moers		
smokers	0.03*	2.09
non-smokers	0.28	–
Munich		
smokers	0.01*	4.96
non-smokers	0.10	–
Saarbrücken		
smokers	0.35	–
non-smokers	0.049*	11.09

* statistically significant: $p < 0.05$

Appendix

A. PAVA procedure

Focusing on binary response the PAVA for increasing trend can be described as follows: Consider the situation of *k* dose groups where the dose d_i , $i = 1,...k$ is in increasing order and the endpoint is the probability p_i of an event (here: chronic bronchitic reaction). We wish to have p_i in non decreasing order, given that $d_i \leq d_{i+1}$. If there is somewhere a violator such that $d_i > d_{i+1}$ for some *i*, then the isotonic estimator of both values is needed to be assessed. That is provided by their weighted mean $p^*_{i,i+1} = \frac{p_i n_i + p_{i+1} n_{i+1}}{n_i + n_{i+1}}$. Now the observations *i,i+1* form a block. This process is repeated using the updated probabilities until an isotonic set of p^*_i is obtained. The algorithm assuming decreasing trend is similar.

B. Test for trend

This test is known as the isotonic likelihood ratio test. We define the following hypothesis:

$H_0 : p_1 = p_2 = ... = p_k = p_0$ against the alternative

$H_1 : p_1 \leq p_2 \leq ... \leq p_k$ with at least one strict inequality.

Then the isotonic likelihood ratio test has the following form:

$$T_{01} = 2\sum_{i=1}^{k}[n_i p_i \ln(\frac{p^*_i}{p_0}) + n_i(1 - p_i)\ln(\frac{1 - p^*_i}{1 - p_0})]$$

This test follows asymptotically a weighted chi-square distribution. However, this approximation does not always hold. Thus a conditional permutations test is proposed.

Based on the observed number of patients per time and dust and the total number of events, a large number of randomizations is analyzed. Each worker is characterized by four elements (d_i, t_i, s_i, St_i) with t_i denoting time, d_i dust, s_i smoking habits, and St_i disease status (0 absence, 1 occurrence). For the permutation test dust d_i is considered independent from (t_i, s_i, St_i) so that the events occur in random allocation. Within each permutation H_0 is considered, the isotonic model is obtained and test T_{01} is assessed. If the observed test value exceeds the 95th percentile, the observed allocation of events can not be explained by chance and H_0 is to be rejected. The exact p-value is estimated as the probability that the result of a permutation is equal or greater to the observed.

References

Baccetti P, (1989) Additive Isotonic Models. JASA 84:289-294

Chuang-Stein C, Agresti A (1997) Tutorial in Biostatistics: A review of tests for detecting a monotone dose-response relationship with ordinal response data, Stat Med 16: 2599-2618

DFG (1978) Chronic Bronchitis and Occupational Dust Exposure. Harald Boldt Verlag, Boppard

DFG (2002) List of MAK and BAT Values 2002. Wiley-VCH, Weinheim

Greim H et al. (2002) Official Answer of the German MAK-commission to the review by McLaughlin et al. Int Arch Occ Env Health

Morton-Jones T, Diggle P, Parker L (2000) Additive Isotonic Regression Models in Epidemiology. Stat Med 19: 849-860

Robertson T, Wright FT, Dykstra RL (1988) Order Restricted Statistical Inference. J. Wiley, New York

Salanti G, Ulm K (2001) Multidimensional Isotonic Regression and Estimation of the Threshold Value. Discussion Paper No. 234, Sonderforschungsbereich 386 der Ludwigs-Maximilian-Universität München

Ulm K (1991) A Statistical Method for Assessing a Threshold in Epidemiological Studies. Stat Med 10: 341-349

Ulm K, Dannegger F, Spanier M (1996) Bericht über die Auswertung der Daten der Studie “Chronische Bronchitis”. Discussion Paper No. 39, Sonderforschungsbereich 386 der Ludwigs-Maximilian-Universität München

Ulm K (1999) Nonparametric Analysis of Dose-Response Relationships. Ann NY Acad Sci 895: 223-231