Statistical tools for analysing the data obtained from repeated dose toxicity studies with rodents: A comparison of the statistical tools used in Japan with that of used in other countries

Author Details

Katsumi Kobayashi  
National Institute of Technology and Evaluation (NITE), 2-49-10 Nishihara, Shibuya-ku, Tokyo - 151 0066, Japan  
e-mail: kobayashi@beige.plala.or.jp

K. Sadasivan Pillai  
Frontier Lifeline, International Centre for Cardio-Thoracic and Vascular Diseases, Chennai - 600 101, India

Soma Guhatakurta  
Frontier Lifeline, International Centre for Cardio-Thoracic and Vascular Diseases, Chennai - 600 101, India

K.M. Cherian  
Frontier Lifeline, International Centre for Cardio-Thoracic and Vascular Diseases, Chennai - 600 101, India

Mariko Ohnishi  
National Institute of Technology and Evaluation (NITE), 2-49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan

Abstract

In the present study, an attempt was made to compare the statistical tools used for analysing the data of repeated dose toxicity studies with rodents conducted in 45 countries, with that of Japan. The study revealed that there was no congruence among the countries in the use of statistical tools for analysing the data obtained from the above studies. For example, to analyse the data obtained from repeated dose toxicity studies with rodents, Scheffé’s multiple range and Dunnett type (joint type Dunnett) tests are commonly used in Japan, but in other countries use of these statistical tools is not so common. However, statistical techniques used for testing the above data for homogeneity of variance and inter-group comparisons do not differ much between Japan and other countries. In Japan, the data are generally not tested for normality and the same is true with the most of the countries investigated. In the present investigation, out of 127 studies examined, data of only 6 studies were analysed for both homogeneity of variance and normal distribution. For examining homogeneity of variance, we propose Levene’s test, since the commonly used Bartlett’s test may show heterogeneity in variance in all the groups, if a slight heterogeneity in variance is seen any one of the groups. We suggest the data may be examined for both homogeneity of variance and normal distribution. For the data of the groups that do not show heterogeneity of variance, to find the significant difference among the groups, we recommend Dunnett’s test, and for those show heterogeneity of variance, we recommend Steel’s test.

Key words

Statistical tools, Toxicity study, Rodents, Cluster analysis, Dunnett’s test, Steel’s test

Introduction

It is a regulatory requirement in most of the countries that toxicity studies are conducted in animal models as per appropriate guidelines for registering industrial chemicals with the respective Government agencies of the countries. Though most of the regulatory guidelines, for example, OECD (1995), EPA (2000) and FDA (2003) give sufficient information on the conduct of repeated dose toxicity studies with rodents, none of them gives a clear picture on the statistical tools to be used for analysing the data obtained from these studies. However, it is mentioned in these guidelines that the statistical methods should be selected during the design of the study. Selection of a non-appropriate statistical tool during the design of the study or using a different statistical tool from that mentioned in the study plan with improper justification at the end of the study may lead to
Table 1: Statistical methods used in various countries to analyse the data obtained from repeated dose toxicity studies with rodents

|------------------------------|-----------------------------------------------------------------------------------------------------------------------|
Comparison of statistical tools used in Japan with other countries

Iran

Israel

Israel

Israel

Italy

Japan

Japan

Japan

Japan

Japan

Japan

Japan

Japan

Japan

Japan

Japan

Japan

Japan

Japan

Japan

Jordan

Jordan

Korea

Korea

Korea

Malaysia

Malaysia

Mexico

Mexico

Mexico

Netherlands

Netherlands

Nigeria

Nigeria

Nigeria

Nigeria

Nigeria

Pakistan

Pakistan

Philippines

Poland

Poland

Poland

Poland

Poland

Poland

Portugal
misinterpretation of the data. This may have tremendous negative impact in assessing the safety of the chemical, as the regulatory bodies heavily rely on these data for assessing the safety.

We made an attempt to compare the statistical tools used in 45 countries for analysing the quantitative data obtained from 127 repeated dose toxicity studies with rodents, and found that the tools we used are not always the same. The parametric data were analyzed by Dunnett's test and the nonparametric data were analyzed by Dunn's test, as recommended by the U.S.A. (2007) guidelines. This may explain why the results of the toxicity studies differ among different countries. The tools used for statistical analysis are important in assessing the safety of the chemical, as the regulatory bodies heavily rely on these data for assessing the safety.

We made an attempt to compare the statistical tools used in 45 countries for analysing the quantitative data obtained from 127 repeated dose toxicity studies with rodents, and found that the tools we used are not always the same. The parametric data were analyzed by Dunnett's test and the nonparametric data were analyzed by Dunn's test, as recommended by the U.S.A. (2007) guidelines. This may explain why the results of the toxicity studies differ among different countries. The tools used for statistical analysis are important in assessing the safety of the chemical, as the regulatory bodies heavily rely on these data for assessing the safety.
Comparison of statistical tools used in Japan with other countries

Fig. 1: Classification of statistical analysis methods by cluster analysis. Note: 11 studies in cluster 1 (red), 2 studies in cluster 2 (orange), 109 studies in cluster 3 (green) and 5 studies in cluster 4 (blue)
used were not similar. For example, to analyse the data obtained from repeated dose toxicity studies with rodents, Scheffé’s multiple range parametric and non-parametric and Dunnett type (joint type Dunnett) tests are commonly used in Japan (Sakaki et al., 2000), but in the other countries use of these tools is not so common. However, statistical techniques used for testing the data obtained from these studies for homogeneity of variance and inter-group comparisons do not differ much between Japan and the other countries. In most of the countries investigated, including Japan, the data are not tested for normality. In Japan, the analysis is usually carried out as per a decision tree (Hamada et al., 1998; Kobayashi, 2000; Kobayashi et al., 2000).

The data of the repeated dose toxicity studies with rodents for the present investigation are from 45 countries and were obtained from internet.

**Investigational materials and analytical method:** Statistical methods used in the 45 countries to analyse the data obtained from repeated dose toxicity studies with rodents are given in Table 1. Based on the statistical tools used, these studies were grouped in 4 clusters as given in Table 2 and were subjected to cluster analysis (SAS JMP, Ver. 5, USA). For the cluster analysis, an input of ‘0’ was given, when a statistical tool was not used and ‘1’ was given, when it was used.

**Results and Discussion**

The classification of statistical tools used in the 45 countries for analyzing data obtained from repeated dose toxicity studies with rodents by cluster analysis is given in Fig. 1. As per the analysis, 11 studies are grouped in cluster 1, 2 in cluster 2, 109 in cluster 3 and 6 studies are grouped in cluster 4. The detection power of statistical tools grouped in cluster 1 for finding significant difference among the groups is extremely low. If the variance of the groups is unequal, using the statistical tools of this cluster may not show a significant difference in the low dose group. The statistical tools of cluster 2 is close to cluster 1, hence the detection power of this cluster is similar to that of cluster 1. If the number of animals is different in the groups, which is usually seen in repeated dose toxicity studies, the detection power of the statistical tools of this cluster for finding a significant difference is further decreased. The statistical tools of cluster 3, which has high detection power, is commonly used in most of the countries. In cluster 4, statistical tools having high detection power were used to examine both homogeneity of variance and normality.

Seven studies from Japan are grouped in cluster 1 of 11 analytical tools, 2 are grouped in cluster 2 of 2 analytical tools and 6 are grouped in cluster 3 of 109 analytical tools. No study from Japan is placed in cluster 4.

Bartlett’s test was used to examine homogeneity of variance in studies conducted in most of the countries. However, 6 studies used Levene’s test to examine homogeneity of variance, which has less power compared to Bartlett’s test. Shapiro-Wilk and Kolmogorov-Smirnov tests were used in two studies each. Interestingly, statistical tool used for post hoc comparison was not mentioned in 14 studies (Table 3).

The number of animals in the group can greatly influence outcome of the statistical analysis of the study. It is also common to encounter mortality in repeated dose toxicity studies, which results in difference in number of animals among the groups. In such situation, the selected statistical tool may have low power for detecting a significant difference, hence cannot bring out biologically relevant information. Hence, the number of animals to be used in a group in repeated dose toxicity studies may be decided taking into consideration of the death that could occur in such studies. Bartlett’s test is a very sensitive test for testing homogeneity of variance of the data and was used in most of the countries investigated. A slight heterogeneity in variance of the data in one group may result in heterogeneity in variance in the data of all the groups by Bartlett’s test, compelling the data to be subjected to a less sensitive non-parametric test. Therefore, for testing homogeneity of variance, Levene’s test, which has low sensitivity (Kobayashi et al., 1999) may be more appropriate than the Bartlett’s test. Present investigation reveals that the data were examined only in 6 studies for both normality and homogeneity of variance. Ideally, the data may be examined for both normality and homogeneity of variance (Kobayashi et al., 2008). We suggest Levene’s test for testing homogeneity of variance of the data. If the homogeneity of the variance of the groups are not statistically different, we recommend Dunnett’s test, and Steel’s test, if it is different.

**References**


Kobayashi, K., K.S. Pillai, M. Suzuki and J. Wang: Do we need to examine the quantitative data obtained from toxicity studies for both normality and homogeneity of variance. J. Environ. Biol., 29, 47-52 (2008).