How to choose an appropriate experimental design type (Part 1)

Liang-ping Hu, Xiao-lei Bao
Consulting Center of Biomedical Statistics, Academy of Military Medical Sciences, Beijing 100850, China

ABSTRACT: How to choose an appropriate experimental design type to arrange research factors and their levels is an important issue in experimental research. Choosing an appropriate design type is directly related to the accuracy and reliability of the research result. When confronting a practical issue, how can researchers choose the most appropriate design type to arrange the experiment based on research objective and specified situation? This article mainly introduces the related contents of the single-group design and the paired design through practical examples.

KEYWORDS: statistics, medical; research design; matched-pair analysis; factor analysis; statistical

Experimental design type is one of the core contents (the three elements, the four principles and the design type) of an experimental design. What is the experimental design type? It is a frame to arrange the experimental factors, block factors (the important non-experimental factors) and their levels. Before an experiment starts, a researcher should choose an appropriate design type based on the research objective and the practical situation; after the experiment, when the data are obtained, it is necessary to correctly judge the design type in order to correctly express and analyze the data. Based on the analysis method for quantitative data, the experimental design can be divided into two types: one is the design type using the deviation analysis, which is also known as non-regression design; the other is the design type using the regression analysis, which is abbreviated as the regression design. In this article, we will introduce the related contents of two non-regression designs, namely, the single-group design and the paired design.

1 The single-group design

1.1 Example 1 The average value of acetylcholinesterase (AchE) of healthy people is known as 1.44 U. A researcher expected to know whether there is a difference of the AchE value between healthy people and patients suffering from chronic bronchitis. Which design type should he choose to

<table>
<thead>
<tr>
<th>Open Access 开放获取</th>
<th>Submission Guide 投稿指南</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOI: 10.3736/jcim20120603</td>
<td>Journal of Chinese Integrative Medicine (JCIM) or Zhong Xi Yi Jie He Xue Bao is an international, peer-reviewed, open-access journal for the study of complementary and alternative medicine or integrative medicine from all regions of the world. JCIM is indexed in PubMed and Directory of Open Access Journals (DOAJ). JCIM is a member journal of CrossRef. Articles published in JCIM have maximum exposure to the international scholarly community. Submit your manuscript here: <a href="http://mc03.manuscriptcentral.com/jcim-en">http://mc03.manuscriptcentral.com/jcim-en</a> (for manuscripts written in English) <a href="http://mc03.manuscriptcentral.com/jcim-cn">http://mc03.manuscriptcentral.com/jcim-cn</a> (for manuscripts written in Chinese)</td>
</tr>
<tr>
<td><a href="http://www.jcimjournal.com">http://www.jcimjournal.com</a></td>
<td>- No submission and page charges for manuscripts written in English</td>
</tr>
<tr>
<td>Hu LP, Bao XL, How to choose an appropriate experimental design type (Part 1). J Chin Integr Med. 2012; 10(6): 615-618.</td>
<td>- Quick decision and rapid publication</td>
</tr>
<tr>
<td>胡良平，鲍晓蕾，如何选择合适的试验设计方案（一）, 中西医结合学报。2012; 10 (6): 615-618.</td>
<td>Send your postal address by e-mail to <a href="mailto:jcim@163.com">jcim@163.com</a>, we will send you a complimentary print issue upon receipt.</td>
</tr>
<tr>
<td>Received April 22, 2012; accepted April 25, 2012; published online June 15, 2011.</td>
<td>ISSN 1672-1977. Published by JCIM Press, Shanghai, China.</td>
</tr>
</tbody>
</table>
carry out the experiment[1]? 

In Example 1, the average value (the standard value) of AchE of healthy people is given. The research objective is to examine whether there is a difference of the AchE value between healthy people and patients suffering from chronic bronchitis. Therefore, the researcher needs to randomly select a group of people suffering from chronic bronchitis as a random sample, measure their AchE value and compare the result with the average value of healthy people. Therefore, example 1 is actually an issue of the single-group design.

1.2 Definition of the single-group design The single-group design is to measure the value of a quantitative index (or multiple quantitative indexes) with respect to a random sample from the same population. If there are \( m \) \((m \geq 1)\) quantitative indexes, there will be \( m \) groups of measurements. In order to conduct hypothesis testing to the obtained measurements, \( m \) standard values or \( m \) population means are required to be given. Above is the main idea of the single-group design. When there is only one quantitative index, it is called the univariate quantitative data with the single-group design; when there are \( m \) \((m \geq 1)\) quantitative indexes, it is called the \( m \)-variate quantitative data with the single-group design[1].

1.3 Characteristics of the single-group design The single-group design only involves one level of an experimental factor and the research subjects are not divided into groups by any other experimental factors or block factors (usually called as the important non-experimental factors). The researcher randomly selects a group of research subjects from the population which is expected to be analyzed and measures the values of \( m \) \((m \geq 1)\) quantitative indexes. It is recommended to estimate the needed sample size following the sample size formulas using the given premises. The premise to conduct the hypothesis testing to quantitative data with the single-group design is that the population mean or the standard value of the quantitative index is given. Of course, statistical inference can also be adopted for univariate quantitative data with the single-group design. To be specific, statistical inference of the single-group design refers to test whether the quantitative data is from a population with a certain distribution (for example, the normal distribution, the log-normal distribution) and to estimate the confidence interval, the tolerance interval and the prediction interval.

1.4 Scope of application When the researcher only needs to examine whether there is a statistically significant difference between the given standards (or the theoretical values) and the population averages of \( m \) \((m \geq 1)\) indexes measured from a group of homogenous research subjects under specific experimental conditions, the single-group design can be adopted. Pay attention to the homogeneity of the research subjects. To be specific, the difference of the observed index is only caused by the individual difference rather than other experimental factors (for instance, type of medicine, dose, duration) or important non-experimental factors. For example, gender, blood type, age, state of illness and course of disease all may influence the result. The group of subjects then should be with the same gender, age, blood type, state of illness and course of disease. In terms of research adopting the single-group design, the accuracy of the result mainly depends on the control of the homogeneity of the research subjects.

1.5 Implementation How to implement example 1? Below are the four key points. First, correctly define the healthy people and their population, which means that the research subjects should be people who at least do not have diseases that may influence the value of AchE; second, based on the result of the pilot test and some special requirements, estimate the minimum sample size \( n \) according to the method of sample size estimation when doing hypothesis testing for quantitative data with the single-group design; third, randomly pick up \( n \) research subjects by a random sampling method (for example, the complete random sampling, the stratified random sampling) from a specified
population; fourth, measure the AchE value of the n research subjects by the stipulated standard method.

2 The paired design

2.1 Example 2 A researcher wanted to perform the experimental treatment of acetyl salicylic acid (ASA) on rats using the platelet activation model. Thromboxane 2 (TXB₂) in plasma is used as the quantitative index. The researcher wanted to know whether there was a difference before and after treatment in the TXB₂ (ng/L) of plasma.

In example 2, the researcher expected to know whether there was a difference before and after treatment in terms of the TXB₂ (ng/L) of plasma, therefore, he may carry out the experiment following the below steps. Step 1, randomly choose a group of rats; step 2, measure the value of the plasma TXB₂ of the rats; step 3, perform the ASA treatment on the rats; step 4, measure the value of the plasma TXB₂ of the rats after treatment; step 5, compare the values of the plasma TXB₂ before and after treatment and draw the conclusion. Therefore, the most appropriate design type for example 2 is the self-paired design.

2.2 Example 3 A researcher expected to evaluate the nutritional value of two kinds of pig forage. How should he conduct the experiment?

The research objective of Example 3 is to evaluate the nutritional value of two kinds of pig forage; therefore, the following three plans may be adopted. Plan A: choose 16 piglets as the research subjects and equally divide them into two groups by complete randomization. The piglets in one group are fed by forage A while the piglets in the other group are fed by forage B. The experiment lasts for 30 d. Measure the increased weight of the two groups of piglets for comparison. Plan B: pair the piglets by gender and weight. Two piglets with the same gender and similar weight form a pair. In the total of 8 pairs, randomly divide the two piglets in a pair into two forage groups. The experiment lasts for 30 d. Measure the increased weight of the two groups of piglets for comparison. Plan C: choose 8 litters of piglets, and then select two piglets from each litter with the same gender and similar weight for experiment. Randomly divide the two piglets from the same litter into the two forage groups. The experiment lasts for 30 d. Measure the increased weight of the piglets for comparison.

All the above three plans can answer the question that the researcher posed. However, the accuracy of the results is different. In plan A, the two groups of research subjects are divided by complete randomization without consideration of the important non-experimental factors like gender and weight, therefore, plan A has a poor balance on gender and weight. In plan B, two piglets with the same gender and similar weight are paired by the researcher, which balances the influence of gender and weight on the result. However, it still can not eliminate the influence of the genetic factor on the result. In Plan C, the researcher selects two piglets with the same gender and similar weight from the same litter, which not only considers the influence of gender and weight on the result, but also eliminates the influence of litter (the genetic factor) which is also an important non-experimental factor on the result. Therefore, in the above three plans, plan C is the best; plan B is better than plan A. Besides, based on plan C, if the intake of every meal of each piglet is measured, the average intake of each piglet in the experimental period can be calculated, which can be used as the covariate when doing data analysis. Thus, the reliability of the conclusion will be higher.

Examples 2 and 3 reflect the three different types of the paired design. Example 2 involves the self-paired design. Plan A of Example 3 involves the design of one factor with two levels (which will be introduced in detail in the next article); plan B of Example 3 belongs to the paired design of subjects with similar conditions; plan C involves the twin-paired design.

2.3 Definition of the paired design The paired design is one of the single-factor designs, and is defined as the design that matches every two subjects by certain conditions, or matches every two data of the same quantitative index observed from the same subject.

2.4 Forms of the paired design The paired design has three forms: (1) the self-paired design, of which every two data of each pair come from the same subject; (2) the twin-paired design, of which every two data of each pair are observed from two subjects with the same parental generation; (3) the paired design of subjects with similar conditions, of which every two data of each pair are observed from two subjects with similar conditions but not with the same parental generation.

2.5 Characteristics of the paired design The paired design has double identities, which means that it belongs to the two-factor design from the point of view of design type while it belongs to the single-factor design in terms of data analysis. First, in terms of design type, the paired design belongs to the repeated-measure single-factor design (the self-paired design) or the random block design (the twin-paired design and the paired design of subjects with similar conditions), both of which belong to the two-factor design that is not able to examine the interactions between the factors. Second, in terms of data analysis, the paired design belongs to the single-factor design because when performing data analysis, the two data of a pair are subtracted in order to get the deviation. That is to say, the analysis of the paired design is transferred into the single-group design with the standard value as 0.
In the self-paired design, the research subjects are not divided by any other experimental factors and the paired condition is the subjects themselves (subjects before and after treatment or two symmetric parts of the body). In the twin-paired design, the genetic factor is used as the paired condition. The two subjects with the same parental generation are divided into two groups. In the paired design of subjects with similar conditions, the subjects are paired by some important non-experimental factors but the genetic factor. For example, the research subjects are paired only by weight (an important non-experimental factor), or the subjects are paired by gender, age, weight and physical condition (a comprehensive factor of more than one important non-experimental factor). Then the two subjects of a pair are randomly divided into two groups[1].

2.6 Scope of application The paired design is adopted when the researcher wants to examine the influence of two levels of an experimental factor on the result. When it is feasible and is strictly performed, the result will be with high accuracy.

2.7 Implementation The paired design has three forms, each of which has its own scope of application. When there is only a short time interval between two treatments and one of them is blank, the self-paired design is the most rigorous with the highest reliability. However, if neither of the two treatments is blank and the observed quantitative index is largely influenced by heredity, apparently, the self-paired design is not feasible. In that case, the twin-paired design may be adopted if possible; otherwise the paired design of subjects with similar conditions can be used. The paired design of subjects with similar conditions should be used only when all the important non-experimental factors that have a large influence on the quantitative index are considered and are strictly matched; otherwise, it is not recommended to use the paired design of subjects with similar conditions. In this case, the design of one factor with two levels should be used.

The paired condition must be the only compound reflection of the important non-experimental factors in the research. For example, assume that gender, age and weight all have influence on the result. The paired condition should be the compound result of the three important non-experimental factors, which means that the two subjects in a pair must have the same gender, similar age and similar weight. The stricter the paired condition is, the better the influence of the important non-experimental factors on the result will be balanced, and thus the higher the quality of the paired design will be.

Based on the result of the pilot test and some special requirements, estimate the minimum sample size $n$ according to the sample size estimation method when doing hypothesis testing for quantitative data with the paired design. Then randomly pick up $n$ research subjects by a random sampling method (for example, the complete random sampling, the stratified random sampling, etc) from a specified population. Finally measure the value of the observed quantitative index of the $n$ research subjects by the stipulated standard method.

REFERENCES


胡良平．统计学三型理论在实验设计中的应用．北京：人民军医出版社．2006：44-52．


胡良平．面向问题的统计学——(1) 科研设计与统计基础．北京：人民卫生出版社．2012：166-168．

如何选择合适的试验设计类型(一)

胡良平，鲍晓蕾

军事医学科学院生物医学统计学咨询中心，北京 100850

摘要：如何选择试验设计类型来安排因素及其水平是科学研究的一个重要内容。选择合适的试验设计类型直接关系到研究结果的准确性、科学性和可信度。面对一个实际问题，研究者该如何根据研究目的及具体情况进行最合适的试验设计类型来安排试验呢？本文将结合实际问题，重点介绍单组设计和配对设计的相关内容和具体实施。

关键词：统计学，医学；研究设计；配对分析；因素分析，统计学