**EVIDENCE-BASED PRACTICE**

**SUZANNE M. MAHON, RN, DNSC, AOCN®, APNG**

**ASSOCIATE EDITOR**

# Reading a Research Article

**Part II: Parametric and Nonparametric Statistics**

*Dana Oliver, MT(ASCP), MPH, and Suzanne M. Mahon, RN, DNSc, AOCN®, APNG*

This is the second in a series of articles to help nurses use and understand statistics. The purpose of the series is to assist nurses in critically reviewing published studies and

Data type

for breast cancer). Distributions that deviate from normal are referred to as skewed distri- butions. Skewed distributions can be positive or negative (see Figure 2). Clinically, a person

implementing the ﬁndings of research into clinical practice. The ﬁrst article addressed

Categorical

Continuous

would not expect to encounter a bell-shaped distribution of a pathologic tumor size. (The

basic statistical considerations and types of variables (Oliver & Mahon, 2005). This article will describe appropriate statistical

Nominal

Ordinal

Interval

Ratio

distribution is skewed because a tumor must be large enough to be found.)

Because alternatives exist for the choice of

methods to use when summarizing data col-

NONPARAMETRIC

PARAMETRIC

a statistical test, why don’t researchers simply

lected from a research project. use the nonparametric statistical methods that

Additionally, this article will introduce the infamous p value. A p value is a probability that determines whether a difference between two or more treatment types or interventions

Skewed distribution

Unequal variance

Sample size

have less stringent requirements? One reason is because parametric methods are more ef- fective in providing reliable results as long as the rules or assumptions are not grossly

is big enough, or statistically signiﬁcant, to change the current standard of care. How- ever, interpretation of a p value requires that subjects were assigned randomly into the two or more groups being compared. Therefore, a variety of randomization techniques will be described to illustrate the inﬂuence that sam- pling choices might have on the interpretation of p values. Finally, this article will discuss the interpretation of statistical signiﬁcance in terms of clinical signiﬁcance.

## How to Choose the Appropriate Statistical Method

Two categories of statistical analysis will be discussed in this article: parametric and nonparametric (see Figure 1). Researchers have a number of nonparametric alterna- tives to consider in place of the tradition- ally used parametric methods (see Table 1). Nonparametric methods play two primary roles in statistical analysis: They are used to summarize categorical data (i.e., nominal and ordinal level data) and in place of the commonly used parametric methods for continuous level data. Figure 1 provides a guide for statistical method selection. It il- lustrates four of the characteristics that must be taken into account for statistical test deci-

This ﬁgure provides an overview of how under- standing the data type (categorical or continu- ous) ultimately guides the process of selecting the appropriate statistical test.

**FIGURE 1. ALGORITHM FOR STATISTICAL TEST DECISION MAKING**

sion making. The rules or assumptions for use of parametric methods must be met to ensure that reliable conclusions are drawn. The primary rules include identiﬁcation of data type, appropriate sample size, variabil- ity of the results of the data, and shape of the distribution of the data. The additive effect of these four characteristics contributes to the power of the analysis.

Table 1 provides a list of the more com- monly used parametric and nonparametric statistical methods for the assessment of research data. The nonparametric methods can be referred to as distribution-free methods (Pett, 1997). Distribution-free refers to the lack of a normal or bell-shaped curve, which frequently is the case with clinical data (see Figure 2). In many clinical situations, the “normal” distribution of the data is not bell- shaped, as frequently is the case in other situ- ations (e.g., age of females enrolled in a study

violated. Another reason is because the non- parametric equivalent methods are thought to be less powerful than the parametric methods. (Recall that loss of power decreases the prob- ability of detecting a difference when a differ- ence truly exists.) However, statisticians now believe that nonparametric tests are almost as efﬁcient as their corresponding parametric tests (see Table 1).

## Assessment of the Similarity or Difference Between

**Two Groups Using a p Value**

The ﬁrst table presented in most clinical research articles describes the subjects who participated in the clinical trial. Descriptive variables usually include age, race, diagno- sis, and treatment type. In many studies, two or more groups are compared. For example,

*Dana Oliver, MT(ASCP), MPH, is a biostatisti- cian at the Saint Louis University Cancer Center, and Suzanne M. Mahon, RN, DNSc, AOCN ®, APNG, is an assistant clinical professor in the Division of Hematology and Oncology at Saint Louis University, both in Missouri.*

Digital Object Identiﬁer: 10.1188/05.CJON.238-240

**238** APRIL 2005 • VOLUME 9, NUMBER 2 • CLINICAL JOURNAL OF ONCOLOGY NURSING

**TABLE 1. LIST OF PARAMETRIC METHODS AND THEIR NONPARAMETRIC EQUIVALENTS**

dence, a potential for bias existed because the control group tended to be older than the experimental group (see Table 2). Therefore,

subjects assigned to one of two groups for a pilot intervention study (see Table 2). The table involves the comparison of two groups

 knowing whether the difference obtained of patients newly diagnosed with breast

**PARAMETRIC NONPARAMETRIC EQUIVALENT**

from the new treatment was because of the intervention or because the control group was

cancer. The primary objective of the trial was to determine whether an intervention, a

–

Independent sample t test

Paired samples t test Analysis of variance

(ANOVA)

Repeated measures ANOVA

Pearson correlation coefﬁcient

Weibull

Chi-square/Fisher exact

Mann-Whitney U

Wilcoxon signed ranks Kruskal-Wallis

Friedman Spearman’s rho

Kaplan-Meier survival and Cox proportional hazards

older is difﬁcult. The resulting effect from bias could be nonreliable, nonreproducible, dis- torted (confounding), or, even worse, a false interpretation of the outcome of the study.

## How to Prevent Sampling Bias and Confounding Effects

The best method for avoiding sampling bias is randomization. Subjects are selected and assigned into one group or another using a randomization method. The terms “prob- ability sampling” and “nonprobability sam- pling” describe randomization techniques.

support group, would have a positive effect on the overall well-being of patients over the course of treatment. The investigator ini- tially employed a randomization technique to prevent bias. However, some patients were reluctant to accept an assignment to the nonintervention group. Therefore, the patients were (partially) randomly assigned to one of two groups. The scores obtained from quality-of-life and functional assess- ment instruments then were compared at three time points during and after comple- tion of treatment.

The characteristics used to describe the

the primary objective of many clinical on- cology trials is to assess whether a new treat- ment is better than the current standard of care. Typically, the subjects’ characteristics and demographics speciﬁcally important to the study are described in the form of frequency (%) for nominal and ordinal level variables. Measures of central tendency and variability (e.g., median, range, mean, standard deviation) are used to describe continuous level variables.

The ﬁrst step in any comparative analysis is to determine whether the characteristics of the subjects assigned to one of the two (or more) groups being compared is not biased. Bias occurs when the characteristics of one group are different from the other and a valid comparison cannot be made. For example, in Coward’s (2003) article on self-transcen-

Probability sampling implies the use of random selection as opposed to a nonprob- ability or convenient sample selection (Fink, 2003). Probability sampling incorporates four types: random, stratiﬁed random, sys- tematic, and cluster sampling. Nonprobabili- ty sampling techniques include convenience, snowball, quota, and focus groups. Although the inclusion of randomization in the design of a study helps prevent misinterpretation of summarized results, it cannot guarantee that the two groups will be equivalent.

## Interpretation of the p Value: Statistical and Clinical Signiﬁcance

Coward (2003) provided a good example of a demographic table used to describe the

study population included variables of continuous and categorical type. Age, num- ber of years of education completed, and months since diagnosis are continuous level. Continuous level variables usually require parametric testing. Six additional characteristics listed in the table are of categorical type. The categorical variables can be described further as nominal and ordinal level. Ordinal level variables include ﬁnancial and physical health status, and the remaining four categorical variables are nominal level (race, treatment type, religion, living arrangement). Categorical variables frequently require nonparametric statistics.

The ﬁrst statistical method employed by the investigator was an independent samples t test. This parametric method was used to

300

250

200

**Frequency**

150

100

50

300

250

200

**Frequency**

150

100

50

0

22 32 41 50 60 69 78 88

**Age (Years)**

Standard deviation = 13.33

–

X = 56.0

N = 1,207

0

0.7 1.3 1.9 2.5 3.1 3.7 4.3 4.9 5.5 6.1 6.7 7.3 7.9

**Pathologic Tumor Size (cm)**

Standard deviation = 1.00

–

X = 1.7

N = 1,121

The ﬁgure on the left is from a series of data on the ages of women diagnosed with breast cancer and represent the classic bell-shaped curve. The ﬁgure on the right is a series of data on the size of tumor and is skewed. Figures were made from a sample data set in SPSS® (SPSS Inc., Chicago, IL).

**FIGURE 2. COMPARISON OF NORMAL DISTRIBUTION WITH SKEWED DATA**

CLINICAL JOURNAL OF ONCOLOGY NURSING • VOLUME 9, NUMBER 2 • EVIDENCE-BASED PRACTICE **239**

**TABLE 2. STUDY PARTICIPANT CHARACTERISTICS BY GROUP**

**EXPERIMENTAL COMPARISON**

**GROUP (N = 22) GROUP (N = 17)**

subjects could bias or confound (distort) the interpretation of the primary objective of the study (i.e., scores of the surveys). A difference may exist in the results be- cause the older women most likely will be

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **VARIABLE** | **—****X** | **SD** |  | **—****X** | **SD** |  | postmenopausal and the younger subjectsmay or may not be postmenopausal. Issues |
| **Age (years)**a\* | 46.1 | 7.1 |  | 51.8 | 11.4 |  | related to menopause could confound the |
| **Education (years)** | 17.5 | 3.4 |  | 16.1 | 3.4 |  | results. |
| **Months since diagnosis** | 2.9 | 1.9 |  | 3.7 | 3.0 |  | Another statistical method used by Cow-ard (2003) is the chi-square statistic. The |
|  | **N** | **%** |  | **N** | **%** |  | nonparametric method is used to determine |
| **Religion** |  |  |  |  |  |  | whether statistically signiﬁcant differencesexist in categorical variables used to de- |
| Protestant | 13 | 59 | 14 | 82 |  |
| Catholic | 3 | 14 | 1 | 6 | scribe the subjects. Referring once again |
| Jewish | 1 | 5 | 0 | 0 | to Table 2, statistically signiﬁcant differ- |
| None | 5 | 23 | 2 | 12 | ences existed with treatment and living |
| **Race**Caucasian | 20 | 91 | 15 | 88 | arrangements between the experimental and comparison groups. A higher frequency |
| African American | 0 | 0 | 1 | 6 | of subjects had undergone mastectomy and |
| Hispanic | 1 | 5 | 1 | 6 | hormonal therapy in the experimental group |
| Asian**Treatment** | 1 | 5 | 0 | 0 | than those randomized to the comparison |
| Mastectomy and reconstructionb\* | 12 | 55 | 3 | group. Also, a statistically signiﬁcant higher18 percentage of subjects currently lived with |
| Mastectomy | 16 | 73 | 10 | 59 |
| Lumpectomy | 6 | 27 | 6 | 35 their spouses and children. These factors |
| Radiation therapy | 6 | 27 | 6 | 35 may have inﬂuenced the way subjects re- |
| Chemotherapy | 16 | 73 | 11 | 65 sponded to the survey. |
| Hormone therapyb\* | 8 | 36 | 1 | 6 |
| **Financial status** |  |  |  | **Summary** |
| Quite secure | 3 | 14 | 1 | 6 |  |
| Comfortable | 9 | 41 | 8 | 47 | Researchers often try to use a random- |
| Okay | 7 | 32 | 7 | 41 | ization technique in an attempt to reduce |
| Marginal | 1 | 5 | 0 | 0 | bias and ensure that treatment and control |

Poor

**Physical health status**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Very good | 4 | 18 | 5 | 29 |
| Good | 11 | 50 | 7 | 41 researchers might use parametric and |
| Some disability, but doing okay | 7 | 32 | 5 | 29 nonparametric statistics when analyzing |
| **Living arrangement** |  |  |  |  | data and looking for differences between |
| Alone | 5 | 23 | 6 | 35 | groups. Researchers must consider the |
| Spouse or partner | 3 | 14 | 5 | 29 | types of data and choose the tests that are |
| Spouse and childrenb\* | 13 | 59 | 3 | 18 | appropriate for the variable types to draw |
| Other | 1 | 5 | 3 | 18 | appropriate conclusions. The next article in |
|  |  |  |  |  | this series will address comparison of more |

2 9 1 6

groups are as similar as possible. This article has provided an overview of how

a t test

b chi-square

\*p < 0.05

*Note.* From “Facilitation of Self-Transcendence in a Breast Cancer Support Group: II,” by D.D. Coward, 2003, *Oncology Nursing Forum, 30,* p. 295. Copyright 2003 by the Oncology Nursing Society. Reprinted with permission.

than two groups and repeated measures and other design issues.

**Author Contact:** Dana Oliver, MT(ASCP), MPH, can be reached at oliverda@slu.edu, with copy to editor at CJONeditor@jsobel

.com.

determine whether statistically signiﬁcant differences existed in age, number of years of education, and number of months since diagnosis (continuous variables) between the experimental group and the control group. Table 2 demonstrates that the mean age of the control (comparison) group was older than those randomized to the ex- perimental group (p < 0.05). The p < 0.05 can be interpreted as a less than 5% prob- ability of the results being due to chance. Therefore, the researcher can be more than

95% sure that a signiﬁcant difference in age existed between the two groups. More speciﬁcally, the mean age and variability of the comparison group was statistically signiﬁcantly higher than the mean age of the experimental group.

Examination of the results obtained from the independent samples t test found that a statistically signiﬁcant difference existed in age. However, this does not always imply that a clinical or biologic difference ex- ists. A reader must consider whether older

## References

Coward, D.D. (2003). Facilitation of self-tran- scendence in a breast cancer support group: II. *Oncology Nursing Forum, 30,* 291–300*.*

Fink, A. (2003). *How to sample in surveys.* Thou- sand Oaks, CA: Sage.

Oliver, D., & Mahon, S.M. (2005). Reading a re- search article part I: Types of variables. *Clinical Journal of Oncology Nursing, 9,* 110–112.

Pett, M.A. (1997). *Nonparametric statistics in health care research.* Thousand Oaks, CA: Sage.

**240** APRIL 2005 • VOLUME 9, NUMBER 2 • CLINICAL JOURNAL OF ONCOLOGY NURSING

Copyright of Clinical Journal of Oncology Nursing is the property of Oncology Nursing Society and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.