Translating Research Into Practice: Overview of Principles

TERRI S. ARMSTRONG, RN, PHD, ANP-BC, JEANNINE M. BRANT, RN, MS, AOCN®, and PEG ESPER, MSN, RN, APRN-BC, AOCN®

From University of Texas School of Nursing, Houston; Billings Clinic, Billings, Montana; and University of Michigan Comprehensive Cancer Center, Ann Arbor, Michigan.

The authors have no conflicts of interest to disclose.

Correspondence to: Terri S. Armstrong, RN, PhD, ANP-BC, UTHSC-SON, 6901 Bertner Ave., Houston, TX 77030. E-mail: Terri.S.Armstrong@uth.tmc.edu

The outcome of any serious research can only be to make two questions grow where only one grew before.

—Thorstein Veblen

As advanced practitioners in oncology, we are dedicated to using current best practice. Innovations in the prevention, diagnosis, and treatment of cancer, and the management of symptoms and toxicity of therapy are occurring rapidly. The ability to critically appraise published research is an essential skill in determining its application to practice.

The *Translating Research Into Practice* section is designed to present a critical review of current, clinically relevant research related to the care and management of oncology patients. In addition, concepts related to clinical trial design and statistical analysis will be reviewed. The ultimate goal is to provide a venue to critically review the literature and to develop skills that can be applied to review and analyze research with the goal of translating relevant research findings into practice. This first *Translating Research Into Practice* column will provide an overview of research and statistical principles.

Statistics 101

One of the most daunting tasks in critically reviewing research is learning to understand the statistics used in the study. Issues related to sample and sample size, reliability and validity of instruments, use of descriptive and inferential statistics, and statistical significance and probability are all important to fully understand and critique a study. Table 1 provides definitions of common statistical terms. Further explanation related to specific terms and analysis will be provided in subsequent columns in relation to the articles discussed.

STUDY SAMPLE

The primary objectives when accruing patients to a research study are to (1) ensure that the sample selected adequately represents the population being studied and (2) determine that the sample size is large enough to be able to generalize the findings to the population of interest. Both are important considerations when conducting a study and evaluating the results of published research. Sampling error can occur if the study is overly representative of a subset of the population or if the sample is too small.

The author should identify demographics of the study sample and criteria used for inclusion in the study. Pertinent information regarding the sample should also be included—for example, the absence of a subset of patients who would typically be part of the sample (minority representation, etc.). Ade-

quate sample size is based on a power analysis that considers the level of significance, power of the statistical test employed, and the effect size (i.e., degree to which a phenomenon exists; Hinkle, Wiersma, & Jurs, 2003). Some quick general "rules of thumb" exist for determining power. For example, a sample size of at least 30 is required to use parametric statistics (described below) and 10 subjects per variable is required for regression analysis.

RELIABILITY AND VALIDITY

Instruments or measures used in the study should have demonstrated reliability and validity to measure the concept being studied. An instrument needs to accurately measure the concept in question. If it does not, the results of the study, no matter how well designed, can be questioned. Validity is the ability of the instrument to measure

Table 1. Common statistical terms and definitions

Term	Definition
ANOVA (analysis of variance)	Procedures used for comparing differences among two or more groups, rather than testing each pair of means separately. Used to determine whether differences are due to chance.
ANCOVA (analysis of covariance)	Same as above, but used to analyze differences between groups on dependent variables
Confidence interval	Range of numbers in which one would find the mean for a population with a stated degree of probability (typically 95%). The range is bonded by confidence limits.
Confounding variable	Variable whose relationship with other variables distorts the relationship between the other variables
Dependent variable	Variable that is measured for the effect that a treatment has upon it
Factor analysis	Multivariate technique which allows you to look at the underlying structure of a set of data and explore relationships among variables.
Hazard ratio	Method to report how often an event happens in one group compared to another, over time. A hazard ratio of 1 means that there is no difference between the two groups. If the hazard ratio is greater than or less than 1, there is a difference in the groups.
Mean	Average score
Median	Center score in a distribution
Mode	Most common score
Odds ratio	Measure of the odds of an event (e.g., occurrence of cancer, disease progression, or a symptom) happening in one group as compared to the odds of the same event happening in another group
Probability value (*p*)	Probability that the data are the result of chance and not an actual difference (e.g., $p = .05$ means there is a 5% chance that the results occur because of chance alone)
Reliability	Extent to which the same test or measurement device will produce the same results on repeated trials
T-test	Statistical test that is used to evaluate if scores of groups differ on a single variable (e.g., fatigue improved for patients who exercise regularly versus those who do not)
Matched T-test	T-test that is used to compare scores on the same subjects (e.g., evaluate if patients' knowledge of their treatment is improved after an education session)
Variable	Observable or measured characteristics
Nominal variable	Categoric variable that cannot be ordered (e.g., gender)
Ordinal variable	Variable that can be placed in order, but the distances between cannot be determined (e.g., letter grades)
Validity	
Construct validity	Degree to which a study or measure accurately reflects the specific concept the researcher is trying to measure
Convergent validity	Degree to which there is agreement between the theoretical concept and the measurement or two measurement tools designed to measure the same concept
Discriminate validity	Agreement among results of measurement ratings that should be related and lack of agreement among when no relationship should exist

Note: Sources: Bailar & Mosteller (1992); National Cancer Institute (2009).

what it is supposed to measure, whereas reliability refers to the ability of the instrument to consistently and accurately measure the concept that is being studied (Wood, Ross-Kerr, & Brink, 2006). If an established instrument is adapted in any way for use in the current study, validity and reliability must be reestablished (Polit & Beck, 2006).

DESCRIPTIVE AND INFERENTIAL STATISTICS

Data analyses of studies in the oncology population are complex and can be intimidating to evaluate for the advanced practitioner not involved in research. In general, two types of statistics are used: descriptive and inferential. Descriptive statistics are those that define the characteristics of the sample (Redmond & Keenan, 2002). Examples include age, gender, and ethnicity and often employ reporting numbers, percentages, means, or medians related to the characteristic.

Inferential statistics are used to identify differences or relationships between variables and define statistical significance. The use of statistical significance allows the researcher to rule out relationships that may be due to chance. Inferential statistics can be broadly categorized into parametric and nonparametric tests. Parametric tests are considered more predictable for what would happen in the population as a whole.

Parametric tests include the paired t-test, independent samples t-test, one-way analysis of variance (ANOVA), two-way ANOVA, and Pearson's correlation coefficient. Nonparametric tests include Wilcoxon's signed-rank test, the Mann-Whitney test, the Kruskal-Wallis ANOVA, Friedman's ANOVA, Spearman's correlation coefficient, and Kendall's rank correlation (Oliver & Mahon, 2005). In future columns, these types of analyses will be discussed in more detail.

To use a parametric test, the data must conform to specific criteria (parameters). These criteria include the following: Data should be (1) numerical (not categorical or nominal); (2) from an appropriately sized sample (in general, > 30 subjects); and (3) normally distributed (the random sample would fall in the middle of a bell curve for the total population) (Redmond & Keenan, 2002). If these criteria are not met, nonparametric tests should be used.

STATISTICAL SIGNIFICANCE

Clinical significance is the relevance of the findings to the advanced practitioner or clinician. While a result may be statistically significant, it may not have clinical significance that would mandate a change in practice. There is not a single statistical method to determine clinical significance; rather, practice change is often based on the quality of the evidence and the results from a combination of studies conducted in a particular area of research. Quality and degree of evidence are often weighed by professional organizations and other experts through "levels of evidence" that will be presented in a future column. Future columns will also ask experts in the field to comment on the clinical significance of the paper that is presented.

When evaluating a new technique or therapy, a research report will highlight results that are found to be statistically significant. Statistical significance is used to represent the generalizability of the results, i.e., their applicability to the population as a whole (Hinkle, Wiersma, & Jurs, 2002; Redmond & Keenan, 2002). A common critical level used for significance in the literature is $p < .05$, which translates to less than a 1 in 20 or 5% chance that the results occurred due to random error.

Conclusion

The ability to critically review research reports is a necessary skill to attain best practice. This column will provide information that will assist the advanced practitioner in conducting a critical review of published research by summarizing important work from the literature and highlighting the research and statistical techniques used in the study.

References

Bailar, J. C., & Mosteller, F. (1992). *Medical Uses of Statistics* (2nd ed.) NEJM Books; Boston.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied Statistics for the Behavioral Sciences* (5th ed.). Boston: Houghton Mifflin.

National Cancer Institute. (2009). Glossary of Statistical Terms. Obtained on 5/24/2010 from http://www.cancer.gov/statistics/glossary.

Oliver, D., & Mahon, S. M. (2005). Reading a research article part II: parametric and nonparametric statistics. *Clinical Journal of Oncology Nursing, 9*(2), 238–240.

Polit, D., & Beck, C. (2006). *Essentials of Nursing Care: Methods, Appraisal and Utilization* (6th ed.). Philadelphia: Lippincott Williams & Wilkins.

Portney, L. G., & Watkins, M. P. (1993). *Foundations of clinical research: applications to practice*. Stamford: Appleton & Lang.

Redmond, A. C., & Keenan, A. M. (2002). Understanding statistics. Putting p-values into perspective. *J Am Podiatr Med Assoc, 92*(5), 297–305.

Wood, M. J., Ross-Kerr, J. C., & Brink, P. J. (2006). *Basic Steps in Planning Nursing Research: From Question to Proposal* (6th ed.). Sudbury: Jones and Bartlett.