This material is protected by U.S. copyright law. Unauthorized reproduction is prohibited. To purchase quantity reprints, please e-mail [reprints@ons.org](mailto:reprints@ons.org) or to request permission to reproduce multiple copies, please e-mail [pubpermissions@ons.org.](mailto:pubpermissions@ons.org)

|  |
| --- |
| **EVIDENCE-BASED PRACTICE SUZANNE M. MAHON, RN, DNSC, AOCN®, APNG—ASSOCIATE EDITOR** |
| Reading a Research Article Part III: The Data Collection Instrument |
| **Dana Oliver, MT(ASCP), MPH, and Suzanne M. Mahon, RN, DNSc, AOCN®, APNG** |
| This is the third in a series of articles intended to assist oncology nurses with improving their knowledge of statistics. Previous articles have discussed types of variables and the use of parametric and nonparametric statistics (Oliver & Mahon, 2005a, 2005b). With knowledge of statistics, oncology nurses can critically review published articles and make more informed decisions on the most appropriate standards of care for patients with cancer. |

This article will focus on the processes that require attention before data collec- tion is initiated. The processes include the design of data collection tools or in- struments and estimation of the number of subjects required to produce reliable and generalizable results to be used by other clinicians. The importance of designing or evaluating data collection tools during the initial stages of the de- velopment of research protocols cannot be underestimated.

As with the previous two articles in the series, this article is a continuation of the examination of the data analysis for an intervention study for subjects di- agnosed with breast cancer in a support group (Coward, 2003). Coward’s study will be reviewed using the ﬁrst two steps of the six-step process presented in Table

1. Issues related to identiﬁcation of limi- tations also will be discussed, because limitations should be identiﬁed and noted throughout the research process—not just at the end. The primary objective of Coward’s study was to pilot a support group intervention that promotes self- transcendence perspectives in women di- agnosed with breast cancer. The second

Ideally, all researchers should meet with a statistician at least three times (sometimes more often) during the study process. The ﬁrst meeting is to perform a sample-size estimation, often referred to as a power analysis. The second meet- ing is to design a data collection tool or evaluate the strengths and limitations of using an established tool. The third meet- ing and subsequent meetings take place throughout the process of data sum- marization. Researchers should keep in mind that data analysis is a process—not a one-time analysis of data. It requires ongoing discussions between research- ers and statisticians to ensure clarity and understanding of the questions being asked.

# Estimate the Sample Size

The pilot study enrolled 41 subjects, all diagnosed with breast cancer. The study design incorporated two groups: the ex- perimental group (support group partici-

pants, n = 22) and the control group (did not participate in support group activi- ties, n = 17). The author did not report a power analysis or sample-size estimation, which is common. Many trials, especially pilot studies, do not have sufﬁcient infor- mation available to calculate estimated sample sizes required when designing research projects. In fact, one of the key roles of pilot studies is to obtain prelimi- nary information to justify the need (and expense) for larger studies.

# Assess the Appropriateness of the Design of the Data Collection Tool

Data collection tools serve two very important roles. First, they force re- searchers to identify all (or almost all) of the data elements required to address primary and secondary objectives. Sec- ond, researchers and statisticians can determine how data should be collected

objective was to assess whether changes in well-being would occur over time be- tween patients participating in support groups and those not participating.

### Dana Oliver, MT(ASCP), MPH, is a biostatistician at the Saint Louis University Cancer Center and

Suzanne M. Mahon, RN, DNSc, AOCN®, APNG, is a clinical professor in the Division of Hematol- ogy/Oncology at Saint Louis University, both in Missouri.

Digital Object Identiﬁer: 10.1188/06.CJON.423-426

for optimal statistical summarization. “Salvaging” poorly collected data sets often is difﬁcult for statisticians (Fink, 2003a). One common example is when researchers collect the ages of subjects and then realize that birth dates should have been collected instead. Birth dates allow for the creation of additional vari- ables. By collecting birth dates, research- ers then can use the data (in the form of dates) to calculate age at diagnosis, age at completion of treatment, etc. For this reason, statisticians often recommend that data be collected in the most raw and potentially most usable form. Consulta- tion during the development of instru- ments can help prevent the collection of unusable or limited data sets.

After meeting with a statistician, re- searchers should pilot test their data collection tools or survey instruments. Data collection tools are forms used to collect information relative to study participants (e.g., patient character- istics, laboratory values, drug doses, survival status). Survey instruments are forms used to collect information about patients’ quality of life, cognitive well- being, symptom distress, etc. Survey in- struments can be presented to subjects via four formats: self-administration, interview, structured record review (data collection tool), or structured observation (Fink, 2003b). Researchers often are advised to use their tools to collect data elements for 5–10 subjects, then meet with a statistician, summarize the results together, and make appro- priate changes. Teams often reﬁne the tools, spending less time retrieving only 5–10 patients or charts rather than 200 or more.

Survey instruments typically are struc- tured so that multiple questions are used

to represent and glean information about the same objective. The primary objective also is referred to as an index or domain. For example, David Cella provided validat- ed instruments speciﬁc to various types of cancer such as breast cancer (Brady et al., 1997). One of the Functional Assess- ment of Chronic Illness Therapy surveys was developed to address quality of life in patients diagnosed with breast cancer (FACT-B) (Brady et al.). The survey has ﬁve domains (physical well-being, social and family well-being, emotional well-be- ing, functional well-being, and additional concerns). Patients answer six items or questions using a Likert scale that rep- resents and provides information about patients’ physical well-being in response to breast cancer. Because the scale has been validated with other patients with breast cancer, researchers can see how their subjects compare to others with the same diagnosis.

Many researchers elect to use instru- ments that were used previously by other researchers. An advantage of the strat- egy is that information about reliability and validity of the instruments may be established already. Instruments often have been revised many times and are designed to facilitate easy data collec- tion. Often, a norm exists to which ﬁnd- ings can be compared. For example, the FACT-B has been normed with women with stage III and IV breast cancer (Brady et al., 1997). If a researcher chooses to use the FACT-B, using it with a similar population may be most appropriate. If a researcher tries to use the instrument with women with in situ or early-stage breast cancer, the ﬁndings may not be valid.

Disadvantages to using existing instru- ments are that the instruments may not

have been used with the populations to be studied or may not address a particu- lar domain or construct. In most cases, researchers must obtain permission to use instruments, and sometimes a cost is involved. Authors of surveys also should provide instructions for scoring survey items.

Another option is to create speciﬁc instruments to be used in studies. Ben- eﬁts and limitations exist when choos- ing preexisting, validated instruments. Researchers must keep in mind that the validation process of a survey is speciﬁc to the type of subjects it was initially designed to analyze. A survey also is vali- dated and assessed for reliability speciﬁc to the primary and secondary objectives. Therefore, researchers should revalidate tools for the speciﬁc objectives and sub- ject populations of their current studies. Cresting new instruments can be time consuming. When instruments are not used with large numbers of subjects, researchers may have difﬁculty deter- mining the reliability and validity of the instruments.

**Table 1. Six-Step Data Analysis Process**

**STEP**

**EXAMPLE**

1. Estimate the sample size.
2. Design a data collection tool.
3. Perform descriptive statisticsa.
4. Determine signiﬁcant associations and differ- ences.
5. Assess the strength of signiﬁcant associations.
6. Deﬁne the limitations of the study.

Power analysis

Cronbach’s alpha

Frequencies (%), mean, median, and standard deviation

T tests, chi-square, and analysis of variance

Pearson correlation coefﬁcient and risk

Decreased power and bias

a Descriptive statistics not only serve to characterize the data being studied but also provide for an

examination of missing values and outliers.

Reliability addresses whether infor- mation collected is repeatable or can be replicated. Five types of reliability measurement exist: test-retest, intraob- server, interobserver, alternate form, and internal consistency (see Table 2). Validity also has ﬁve types: face, content, criterion: concurrent, criterion: predic- tive, and construct (Litwin, 2003) (see Table 3). Validity addresses whether an instrument measures the construct or question. Basically, validity answers the question, “Am I measuring what I think?” Reliability answers the question, “Is the information gained repeatable?” If a sur- vey instrument is proven to be valid, then it also is reliable. However, an instrument may be proven reliable but not necessar- ily valid.

Coward (2003) used a demographic data collection form that was modiﬁed (based on previous use) to better assess participants’ current treatment status, stage of disease, and developmental stage at diagnosis. She added additional items to identify factors preventing women from participating in support groups. The participants in the breast cancer support group study were given eight surveys to complete at three time points (see Table 4). The survey tools were as- sessed for reliability and validity.

**Table 2. Types of Reliability Measures**

**TYPE OF**

**RELIABILITY PURPOSE**

**EXAMPLE**

**ISSUES OF CONCERN**

Alternate form

The responses to the questions are

reworded or reordered but maintain functional equivalency or continue to ask and answer the same question.

“Yes or no” versus “no or yes”; re-

word the response as follows: one to two times per day versus 12–24 hours per day.

Internal consistency To assess whether a batch or group of

questions represents the same concept (e.g., emotional well-being)

Interobserver

To assess whether two or more raters

agree

To assess whether an individual rates the same issue in a consistent manner

Six questions are used to assess emo-

tional well-being by the Functional Assessment of Chronic Illness Therapy– Breast Cancer survey.

Four different nurses provide an oral assessment scale (OAS) for the same patient.

A nurse provides an OAS for the same patient (e.g., one in the morning and another in the evening).

Postoperative patients measure pain using a scale two days and seven days after surgery.

Provides a mechanism of compensat-

ing for the practice effect, the idea that a rater or respondent recalls or becomes familiar with questions on a survey after repeated use (Litwin, 2003)

Beware of using a collection of ques- tions that do not “ﬁt well” together.

Observers will have to practice assess-

ment so ratings are consistent.

Intraobserver

Beware of the practice effect.

Test-retest

To measure how stable a person’s re-

sponse is by giving the respondent the same survey at two different times

Beware of variables that change over a

short period of time.

**Reliability**

Always note whether surveys have been validated for reliability to ensure that the participants understood the ques- tions (in the way the researchers intended the questions to be understood) and that the questions addressed the stated pri- mary and secondary objectives. One of the statistical methods used to determine internal consistency is Cronbach’s coefﬁ-

cient alpha. The statistic is a measurement of the strength of the internal consistency (or homogeneity) of a set of survey ques- tions. For example, it is an assessment that measures the extent to which items included on a questionnaire focus on a particular domain (e.g., patient satisfac- tion, well-being).

Table 4 gives a list of the eight tools used by Coward (2003) as indicators of whether the support group intervention

was successful. The table also provides the authors and creators of the study tools and instruments. Mentioning au- thors and creators of instruments is im- portant and professionally considerate. Researchers should obtain permission to use instruments for their own studies. When obtaining permission, researchers also might want to take the opportunity to question authors of tools regarding speciﬁc issues they encountered.

**Table 3. Types of Validity Measures**

**TYPE OF**

**VALIDITY PURPOSE**

**EXAMPLE**

**ISSUES OF CONCERN**

Construct

To measure how meaningful a survey

instrument is, usually after many years of experience

To obtain an opinion from a trained individual

Results from the instrument have

been reported in numerous research projects.

Obtain subjective opinions from social workers, oncologists, and nurses about a survey that assesses quality of life after cancer treatment.

Compare the results obtained from a survey developed by a researcher’s own institution with a previously

validated survey measuring the same indexes.

Use oral assessment scale scores for prediction of pain medication doses.

Not easily quantiﬁable

Content

Provides a subjective opinion, not an

objective measurement, of the appro- priateness of a survey or question

Criterion:

concurrent

To compare a newly developed survey

with a gold standard (previously vali- dated survey)

A gold standard must be available to

use for comparison.

Criterion:

predictive

Face

To ﬁnd the predictable usefulness of a

score, determined from a survey, with some associated patient outcome

To obtain an opinion of the survey from an untrained individual

Not recommended for longitudinal

clinical studies; the time interval be- tween survey and outcome is too long.

Obtain a subjective opinion of the sur- Not considered a true measurement vey from a roommate or spouse. or assessment of validity; usually per-

formed at a pretest stage

Finally, limitations answer questions of whether study results are generalizable to other subject populations at other cancer centers. One of the unexpected limitations noted by Coward (2003) re- lated to a potential for bias. Many of the participants selected to be in the experi- mental group, suggesting that they were aware of the value of support groups dur- ing treatment for breast cancer.

**Table 4. Study Instruments and Reliabilities**

**RELIABILITY ALPHA**

**INSTRUMENT**

**AUTHORS**

**TIME 1 TIME 2 TIME 3**

a Pearson’s correlation

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Self-Transcendence Scale | Reed, 1991a | 0.87 | 0.83 | 0.84 |
| Purpose-in-Life Test | Crumbaugh & Maholick, 1964 | 0.90 | 0.87 | 0.88 |
| Affect Balance Scale | Bradburn, 1969 | 0.83 | 0.60 | 0.71 |
| Proﬁle of Mood States | McNair et al., 1992 | 0.95 | 0.94 | 0.93 |
| Cognitive Well-Being Scale | Coward, 1990b | 0.49a | 0.76a | 0.68a |
| Symptom Distress Scale | McCorkle & Young, 1978 | 0.90 | 0.83 | 0.93 |
| Karnofsky Performance Status | Karnofsky et al., 1948 | – | – | – |
| Personal Resources Questionnaire | Brandt & Weinert, 1981 | 0.88 | 0.92 | 0.93 |

Once researchers obtain permission to use tools, they must test whether the instruments can be used reliably to assess their particular study subjects. Coward (2003) took the steps to verify that the eight previously validated instruments also were reliable for use with the 41 subjects enrolled in her study who recently were diagnosed with breast cancer. Coward carried the analysis a step further by also validating whether the instruments were consistently reliable over the course of the study (pre- and postintervention). Cronbach’s alpha statistic was used as a measurement of reliability for seven of the eight instruments. Levels of 0.7 or higher generally are accepted as representing good reliability (Litwin, 2003). Most of the reliability coefﬁcients (except one) were higher than 0.7, suggesting that the instruments could be used reliably for measuring whether the intervention was successful for patients newly diagnosed with breast cancer in terms of quality of life. Most of the reliability coefficients were relatively high (higher than 0.8) and remained consistent for the three time periods. However, the coefﬁcients for one of the instruments (Affect Balance Scale) tended to exhibit a decrease in reliability over the course of the study. This may have been related to the effects of treatment.

## Validity

Coward’s (2003) discussion of the va- lidity of the instruments used in the study was not nearly as extensive as the discus- sion of reliability. The author included a statement for some instruments that had no known content and construct validity,

but a speciﬁc discussion of the validity was not provided. Although this is com- mon, it requires interested readers to ﬁnd the primary sources for the instruments to better understand their validity and appropriate use in studies.

# Limitations of the Study

Throughout the process, researchers should consider that no study is perfect. Identiﬁcation of limitations is an ongoing process. Often during construction or selection of survey tools, trade-offs are made to balance the usefulness of tools with the feasibility of studies. Limitations can be put to good use, leading to future considerations and studies. Frequently encountered limitations include low sample size, use of surrogate measures, and a heterogeneic population of study participants (e.g., in terms of diagnosis and treatment regimens). Limitations should be identiﬁed throughout the study and corrected whenever possible.

Limitations always should be docu- mented and reported in manuscripts summarizing study results. Awareness of limitations is important to readers for many reasons. Limitations provide not only a better understanding of why result- ing outcomes happened but also valuable information to people who might want to pursue and conduct similar research projects. If follow-up studies were to be conducted by other researchers, the is- sues, if known, might be preventable.

Limitations also give a more accurate description and understanding of the ﬁndings, especially the unexpected is- sues that arise during the course of study.

# Summary

Parts I through III of this series of ar- ticles gave an introduction to some of the issues requiring attention when perform- ing clinical studies and summarizing the results statistically as well as clinically. Sta- tistical methods have been described for the identiﬁcation of statistically signiﬁcant differences and associations between an experimental group and a control group or over the course of time after an inter- vention. The next article will address the methods used for the determination of the strength or magnitude of these identiﬁed statistically signiﬁcant differences.

**Author Contact:** Dana Oliver, MT(ASCP), MPH, can be reached at [oliverda@slu.edu,](mailto:oliverda@slu.edu) with copy to editor at [CJONEditor@ons.org.](mailto:CJONEditor@ons.org)

## References

Brady, M.J., Cella, D.F., Mo, F., Bonomi, A.E.,

Tulsky, D.S., Lloyd, S.R., et al. (1997). Reliability and validity of the Functional Assessment of Cancer Therapy–Breast Quality-of-Life instrument. *Journal of Clinical Oncology, 15,* 974–986.

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

Fink, A. (2003a). *How to manage, analyze, and interpret survey data* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Fink, A. (2003b). *The survey handbook* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Litwin, M.S. (2003). *How to assess and in- terpret survey psychometrics* (2nd ed.). Thousand Oaks, CA: Sage Publications.

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

Oliver, D., & Mahon, S.M. (2005b). Reading a research article part II: Parametric and nonparametric statistics. *Clinical Jour- nal of Oncology Nursing, 9,* 238–240.

Copy r ight of C l i nical Journal of Oncology Nu rsing is the property of Oncology Nu rsing Society and i ts content may not be copied or emai led to mul ti ple sites or posted to a l istserv wi thout the copy r ight holder 's express w r i tten permission. However. users may pr i nt. download. or emai l articles for i nd i vidual use.