Promoting Advanced Practice Nurse Buy-In For Evidence-Based Practice: An Apn Research Guide

S Northam, J Lakomy

INTRODUCTION

Reading and understanding research is challenging and is often cited as a barrier to research application (1-3). Generally novices require more time to read, understand, and critique studies than experienced advance practice nurses (APNs) (4), but many find the research critique to be a helpful learning activity (5). Lack of familiarity with terms and lack of confidence in ability to understand the findings can make the experience of reading and critiquing research intimidating and frustrating. Few nurses regularly read research; and alarmingly, Wood (6) recently noted that trepidation, antipathy, and apathy toward research have not improved in the past two decades (7). Regular reading to gain insight into best evidence useful for practice is essential if nursing is to meet the goal of improving health care through the implementation of evidence-based practice (8). Thus, strategies to improve the skills of APNs in reading and understanding research can make the experience less overwhelming, time consuming and frustrating. Recent findings have shown that reading research increases critical thinking skills (9), thus benefiting not only patients but also nurses. This article reviews an APN Research Guide devised, used and refined for 12 years by nurses who have evaluated the guide positively. Data are included to support their evaluations. Understanding research is an essential first step towards evidence-based practice (10-12).

This APN Research Guide fills a major gap in critique guides. This guide is short (4 pages), easily printed with space for responses, and contains the first scoring criteria since Duffy’s criteria was published in 1985 (13). The point allocations for this critique were derived by the authors who are experienced clinicians and researchers. While other critique guides exist, the ones in research texts are generally long, presented as text boxes within chapters, and lack space for readers to include notes as they critique.

Polit and Beck (2004) used 138 questions within 18 text boxes to guide readers in evaluating published quantitative research reports. Burns and Grove (2003) included 38 sections with 150 questions: 19 sections with 56 questions in the comprehension guide; 11 sections with 86 questions in the comparison and analysis guide; and 8 questions in the evaluation guide.

The Burns and Grove guide includes more directing questions for evaluating the literature review, separates evaluation of variables from measurement and validity assessment, and includes questions about missing study elements. The reader is also guided through the process of article review and critique steps twice rather than once. Both Polit and Beck (2004) and Burns and Grove (2003) recommend that readers summarize study strengths and weaknesses but do not recommend scoring research studies. The multiple sections displayed as text boxes in several chapters cannot be quickly printed and easily used by nurses in practice.
Recently Daggett, Harbaugh and Collum (2005) published a single page worksheet devised to teach baccalaureate nursing students to critique research by briefly noting parts of the research. The worksheet is utilitarian in its simplicity but is more geared to noting the parts of the research rather than critiquing the study or determining the utility for evidence-based practice. Cutcliffe and Ward (2007) argue that no one best guide to reading and critiquing research would suit all nurses. These authors review various methods of critique, note the strengths and limitations of each, and end with a presentation of a new approach used by the Network for Psychiatric Nursing Research (NPNR) Journal club. The NPNR Journal club approach can be used for both qualitative and quantitative research and includes eleven guidelines. The NPNR approach lacks guidelines in reviewing the scientific method of quantitative research, derives no scoring of the research report but encourages nurses to support a critique by listing strengths and limitations.

In many areas of life from kindergarten to delivering advanced patient care, steps are often created that break the task into small parts. Small parts create the impression the whole task is not overwhelming and provide a mental advantage and organization for completion. The APN Research Guide accomplishes that goal of breaking the task of reading and understanding research into small parts.

Throughout the APN Research Guide, point allocations are noted among the steps of research to enable readers to calculate an overall critique score, which has not been done since Duffy’s critique in 1985 (13). While Duffy’s maximum score was 306, the maximum on this APN Research Guide is 100, a more traditional score. The higher the score, the better the nurse appraises the parts of the research study and evaluates its strengths overall. A rating over 80 would indicate the research has high quality and should be highly considered when the APN is making a clinical decision in the practice setting. A future article will provide readers with an example of a critique with a high score and utility for implementation in evidence-based practice by advanced practice nurses.

THE MAJOR STEPS OF THE RESEARCH PROCESS

THE PROBLEM

Although simple sounding, discerning the research problem can be difficult and may be confused with the purpose. For example, while the purpose of the research may be to identify ways to safely see more patients in a busy clinic setting, the problem may be long wait times to get lab results. Discerning the problem allows for a much more focused way to view how to study, measure, and improve the situation. The problem is the one clearly addressed by the study rather than an expansive area such as efficient health care delivery. Identifying the problem also lets advanced practice APNs see if this research study is relevant to problems or issues in their own sphere of influence and practice.

REVIEW OF THE LITERATURE

Generally the initial narrative of the research article, just after the abstract, lays the groundwork for what is known about the topic. The discussion of existing literature cites prior research and helps the APN understand what is already known about the problem. Some journals restrict the length of articles and thus the discussion of existing literature may be brief. Classics are well known works in an area. For example, if the study is on dreams and their meaning, Freud likely would be included since he wrote extensively on dreams. Beginning readers may be unfamiliar with classics, but this section serves as a reminder that there is classic research on most issues in nursing and allows the APN to make a judgment call as to whether or not the review of the literature is sufficient if changes in practice are being considered.

THEORETICAL FRAMEWORK

Just like a long road trip requires a map, a research study needs an organizational structure or framework, commonly called the theoretical framework. Quantitative research involves deductive reasoning flowing from the framework to a testable hypothesis designed to support or refute theory. Evaluation of the reasoning or conclusions that were reached by the researcher is important. Strong research clearly explains the underlying theory to be tested by the research and sometimes supports the theory with a visual model.

RESEARCH QUESTIONS AND HYPOTHESES

All quantitative research involves either research questions and/or hypotheses, and all research reporting significance used statistical analysis to test hypotheses. Frequently the research questions are noted at the end of the article’s introduction. A research question is the researcher’s underlying probe as to why something is happening (or not happening) in certain situations. For example, the question may be as simple as “Will providing elderly patients with pedometers increase strength and endurance in walking
exercises?” A hypothesis that flows from the theoretical framework is more of a speculation about what will happen if certain situations occur. The researcher has more of a burden with hypotheses which must be supported or not. An example of a hypothesis would be “Providing pedometers to elderly clients will increase the amount and quality of walking by 20% over a six month period.” Beginning readers may find that selecting journals that include a hypothesis heading within the abstract is helpful. If the hypothesis is not clearly stated some sleuthing may be needed. A quick check in the results section may report the hypothesis along with the type of statistics used to test the hypothesis. Knowing the research question or hypothesis will help the APN understand what to look for in the discussion of findings.

**RESEARCH DESIGN**

Quantitative research generally progresses along a continuum. When little is known about a phenomenon, researchers may simply aim to describe it and are thus at the descriptive end of the continuum. When variables are known and surveys can be designed or existing surveys can be used to examine relationships among variables, the design is at the survey point, sometimes also called correlation research. Research in which the researcher directly plans the manipulation of an independent variable and measurement of a dependent variable is at the experimental end of the continuum. This distinction is important as APNs try to decide whether or not to recommend a change in practice based on the appraisal of the evidence. Descriptive and correlation research can show relationships and situations that are similar to those APNs face each day and can offer some alternatives. However, before implementing evidence-based practice, the APN must appraise the evidence for soundness, evaluate the sample size for adequacy, conclude the findings are strong enough to warrant a consideration of changing the ways things are done, and determine that the proposed intervention is congruent with patient preferences. Quasi-experimental means “almost, but not quite, experimental.” When human subjects are studied, sometimes all of the variables cannot be controlled like can be with animal models. Using quasi-experimental does not mean the research is not sound. Whether or not the study is a quasi-experimental design is dependent on the use of randomization and/or a control group. This section of the critique includes brief questions applying common research terms.

**POPULATION**

Research texts generally restrict the population to the group from which the sample was drawn. Research rarely uses findings for the exact same population but rather evidence-based research efforts are geared to moving research to more widespread, yet appropriate, use. This section is designed to encourage the APN to think about the demographics of the population and who could benefit from the findings. For example, a study involving mothers receiving group prenatal care that involves a New England population of 1000 Whites, Blacks, and Hispanics ages 16-42 can be used in other areas of the US but lacks knowledge about Asians. Clearly, research done on one or two ethnic or racial groups can hardly be generalized to all. The challenge is to recognize best evidence and improve practice through the appropriate generalization of research.

**VARIABLES**

In order to determine the usefulness and appropriateness of applying research findings in the clinical setting, the APN must determine what is being tested. The variables must be defined. One researcher might describe “lifting risk” as the chance of dropping a patient on the floor; another researcher might describe the same variable as the chance the APN will get a back injury. Both are essentially right, but the reader must know exactly what is being discussed, tested and recommended; therefore, quantitative research involves both conceptual and operational definitions. The conceptual definitions are the ones mentioned previously; the operational definition is how the researcher measured the variable, such as a patient fall risk scale or a back injury scale. The research consumer must decide if the measurement of the variable makes sense and truly reflects the concept of interest. This discussion pertains to construct validity. APNs can often discern the variables by referring back to the research question and/or hypothesis. The title also generally includes the key variables. The table within the critique provides a place to list the name of the variable (column 1), then note how the variable was measured or operationalized (column 2), and if that method of measuring the variable is valid. More than yes or no answers are necessary and require critical thinking. The key critical thinking is whether the measurement of variables done in the study enables the APN to know about the underlying concepts of interest. Actual surveys are not included in publications but sometimes sample survey items are included. The APN can evaluate the clarity and appropriateness for the population of sample items. The total
number of survey items can also be noted. More items sometimes improve the ability to capture the concept, but certainly expansive, time consuming surveys are likely to undermine completion due to fatigue. Evaluating validity can be difficult, but practice will increase the APN’s skill.

RELIABILITY
The term reliability has the same basic meaning in research as in other areas such as employment, appliances, and automobiles, among other examples. But the myriad of types of reliability, particularly the common internal consistency, takes the term beyond the general understanding. This section in the APN Research Guide was designed to narrow down reliability to the three types: stability, internal consistency, and equivalence. Stability means that similar results are obtained on two separate administrations of the instrument. Homogeneity or internal consistency states that the items on the instrument measure the same trait. Interrater reliability, a type of equivalence, means that two or more independent researchers recorded or observed the events in the same manner. This section also includes a section for APNs to note any actions taken to enhance reliability. Generally this area is a challenging part of completing the guide.

PILOT
Researchers often do a dry run of a study to work out any problems before undertaking a large scale, extensive study. The pilot may simply involve testing surveys using a small group to discern item clarity and the time required to complete them. Pilots are also helpful when effect sizes are unknown and initial data can assist in examining for an effect size to determine the sample size required in a larger study to discern the effect.

SAMPLE
The number of individuals who actually participated and provided data are considered to be the sample and are generally included in every research study. When surveys are used, individuals may elect to answer some items but not others and this occurrence explains why the n (sample size notation) varies within tables or the article narrative. Randomization may occur in either selection or assignment. Random selection begins with a source of individuals and randomly selects a sample from it. Telephone surveys, for example, randomly select a designated sample size from all available telephone numbers. Random assignment places individuals into groups using a method such as a coin toss. There are a variety of methods of random assignment and random selection but neither is haphazard. The goal of randomization is to allow individuals an equal chance of selection or assignment and reduce bias. Randomized controlled designs (RCT) are the classic experimental research design and involve random assignment into groups. Sample demographics, often presented in a table, can be compared to the demographics of the population to evaluate if the sample is representative.

DATA COLLECTION
The actual steps involved in gathering data are generally addressed very briefly within the research article, primarily to save space. The reader may be unable to discern who collected the data or the circumstances of data collection. Acknowledgements at the beginning or end of the article may note the names of individuals who gathered data but all authors may not have actively participated in the data collection. As much information as possible can be included in this section but incomplete insight into the “who, what, when, where and how” is common.

LIMITATIONS
The limitations section primarily pertains to internal and external threats to validity. Discerning threats helps to evaluate rival hypotheses, which are alternate explanations for changes in the dependent variable other than the independent variable. Controls are any actions taken by the researcher to understand or minimize the threats to validity.

PREPARING DATA FOR ANALYSIS
This section was devised as a step before considering the statistical analysis. The variables listed here should directly mirror each variable listed in section 7 of the APN Research Guide where the variables, their measurement and validity were presented. The goal is to think of what the data will yield in terms of a number and whether an average is feasible. Non-parametric statistics are used when no average score is possible or meaningful such as with demographic variables of race or marital status. Parametric statistics are used when an average score is possible and meaningful with the caveat that other assumptions for each parametric statistic must also be met. So by determining whether a meaningful mean is possible from the data, the reader gets a beginning idea of whether non-parametric or parametric statistics will be used in analysis.

DATA ANALYSIS
Many research reports include a variety of statistics from descriptive statistics used to report demographics to
inferential statistics used to test hypotheses. At this point in the critique the APN should refer back to the hypothesis and look for reported significant or non-significant findings. Inferential statistics involve probability testing and report a p value which equates to the likelihood of error. In research with minimal risk, a preset alpha of .05 is generally used and indicates a 5% chance of error, or 5 out of 100 times the results are due to chance. A more conservative preset alpha of .01 is used for research with more risk such as clinical trials involving new medications. Hypothesis testing is done based upon the wording of the hypothesis, the number of groups and the number of variables. Hypotheses which predict a relationship between two continuous variables are tested using Pearson’s Product Moment Correlation. Hypothesis predicting a difference in two groups on a single continuous outcome variable are tested using a parametric t test. Differences in more than two groups on a single continuous variable are tested using ANOVA. When a number of independent variables are used to predict a single continuous outcome variable, multiple regression is used and if the outcome variable is dichotomous, logistic regression is appropriate. Many research studies involve testing of more than one hypothesis and thus the results section includes findings using a variety of statistical tests. The pre-set alpha is the level of error the researcher felt reasonable before beginning the research and is often omitted from the research report. Since a 5% chance of error is common, any reported p values less than .05 that are termed significant (sometimes using an asterisk) imply a preset alpha of 0.05. Sometimes notations under tables help to discern the preset alpha. A single asterisk is conventionally used to denote significance found at the ≤05 level; two asterisks denote significance at ≤.01, and three asterisks denote significance at ≤.001. Significant findings indicate the findings are unlikely due to chance and similar results can be expected from similar populations, thereby facilitating generalization of the findings.

**INTERPRETING THE RESULTS**

The results of research are interpreted based upon the statistical findings, their credibility and sample size. When findings are statistically significant and a hypothesis was tested, the results support the hypothesis. For example, in a study that hypothesized women in group prenatal care experience more maternal stress reduction during pregnancy than standard prenatal care, a p value of <.05 and a lower mean stress level in the group prenatal care group supports the hypothesis. If the results make sense the reader evaluates them as credible. Studies may test more than one hypothesis and the most important one is often reflected in the journal article title. The groups would need to be fairly similar initially on stress levels as well as demographic variables to ascribe the changes to the independent variable, the type of prenatal care. Large samples lend credibility to the findings and small samples undermine credibility. At least 30 people per variable and even group sizes (17) are desired. Skilled clinicians reading in an area of expertise should trust their ability to interpret research findings.

**COMMUNICATION OF FINDINGS**

Internationally, researchers strive to reduce the time from research to implementation into practice. While all research cannot be easily implemented, this section prompts the APN to evaluate who should know about the research and whether or not implementation is appropriate. The actions should clearly be derived from the research. Small, simple interventions such as the use of music to reduce procedural stress may be easily implemented in practice. New procedures, such as the method of caring for wounds, generally require more evidence and policy changes. Research reports often include a section on implications for practice which summarize recommendations for using the study findings.

**ETHICS**

Four areas are included to evaluate the three basic tenets of research ethics: beneficence, respect for human dignity, and justice. APNs are directed to comment on institutional review board (IRB) review, informed consent, risks, benefits, and the qualifications of the researchers. Any reference to grant funding, particularly US government funding from institutions like the National Institute of Health (NIH) or the American Cancer Institute, imply IRB approval since such approval is required at the time of proposal submission. The use of a consent form also implies IRB or human subjects review board (HSRB) approval. Risks are rarely enumerated within research article reports so the reader must consider the risks and weigh them against the benefits. The qualifications of the researchers are determined by noting the educational and professional credential notations after the author names or within a footnote about the authors. Anonymity exists if the researcher never saw the research subject. More often anonymity did not exist within the research study but confidentiality was protected.
INTERPRETING THE CRITIQUING SCORE

Once the APN has completed the APN Research Guide, a score is obtained. If the score is less than 70, no change is recommended for practice. If the score is between 71 – 79, the APN should utilize this research cautiously in the practice setting. If the score is 80 or greater, the APN may more comfortably include the findings in the clinical decision process for evidence-based practice. Scoring on a 100 point scale is common in graded educational projects and the 71-79 scores equate to average work, 80 or greater is good, and 90 or greater is excellent. The cut off scores were derived from the authors’ educational experience and used successfully by clinicians enrolled in graduate education. A pilot study of the scoring was conducted using 4 nursing graduate students enrolled in a research course that read and rated a published primary research report 5 weeks into the course. Inter-rater reliability was calculated as .89 (M = 89.5; SD = 5.26; Std. Error = 2.63) and scoring ranged from 85-97 without any specific training. The nurses were simply asked to read and use the APN Research Guide, rate each section, and then create a sum score. The high inter-rater reliability reflects the good consensus in scoring the research and indicates there was general agreement on the quality of the published research report. The nurses agreed the article was good and amenable to implementation for practice.

EVALUATION OF THE CRITIQUE

This critique guide was devised to help APNs read and evaluate research. The guide has been very positively evaluated and APNs suggested this article as a way to share the tool. To gather some recent evaluation data, an anonymous online evaluation was completed by APNs (N=12), 98% strongly agreed and 8% agreed that the guide assisted their understanding of research. Qualitative comments included: 1) “I enjoyed the critiques--they changed my way of reading professional articles and studies. They made me think a little more critically”; 2) “It has helped me to become more comfortable reviewing research and discussing it with others. Information I have seen before now makes more sense to me”; 3) “Even though I started out fearing research, writing the critiques made me understand the process so much better. Now, I look for articles to see if they are as great as they say they are. What a great learning experience.”

SUMMARY

Quantitative research involves a deductive process outlined within this critique. The goal is to break the process down into small parts with questions to guide APNs in doing critical appraisal of the evidence. Even when a section or two are challenging, the overall ease of completion fosters the understanding and confidence that are essential for APNs striving to improve their skills in reading and using best evidence in practice.

ACKNOWLEDGEMENTS

Special thanks to Dr. Lynn Wieck for editorial input and Dr. Cheryl Beck for review and input.

Figure 1

APN RESEARCH GUIDE

The major steps in the research process are presented. Use this guide to make notes and then assign a percentage (%) of the point allocations based upon your assessment of each step.

1. Identify the problem. Briefly state it in your own words: (6 points)
2. Review of the literature: (6 points)
   At the beginning of the article, did the author review other research studies that have been done on this topic?
   Did it appear that “classic” had been reviewed?
   Were the majority of references recent? (≤5 years’ total # references)
3. Theoretical/Conceptual Framework: (6 points)
   Is it clear to you how this research fits into nursing?
   Does the literature review lay a clear, solid groundwork for the study?
4. Formulation of research question(s) or hypothesis(es): (6 points)
   Are there research questions? If yes, write the hypothesis(es) here:
   Are there hypotheses? If yes, write the hypothesis(es) here:
   Are they null, non-directional or one-way?
(Note: All quasi-experimental or experimental research MUST have hypotheses, even if you need to figure out what the hypotheses are based upon the stated problem.)
5. Selecting the research design: (6 points)
   Underline the terms that apply to the research design of this study:
   1. Cross-sectional OR Longitudinal
   2. Descriptive OR Survey OR Quasi-Exp. OR Experimental
Figure 2

6. Specifying the population. (6 points)
   What is the population? (Include some demographics in your answer)
   (Note: The population is NOT the same as the sample. The sample is drawn from the population. So this answer should not include a number)

7. Operationalizing and measuring the research variables: (8 points)
   List each variable here and explain how the researcher(s) measured (operationalized) each variable. Then critique the validity of each variable.

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>HOW THE VARIABLE WAS MEASURED</th>
<th>VALIDITY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What is the independent variable(s)?

What is the dependent variable(s)?

8. Reliability: (All of the following questions pertain to reliability) (8 points)
   a. What Cronbach’s Alpha was reported for the instrument? (= internal consistency). What do you think of that Cronbach’s Alpha?
   b. Was a test–retest reported? (=Stability)
   c. Was any inter-rater reported? (= Consistency)
      If yes, what was the % and what do you think of that?
   d. If another researcher replicated this study is there a good chance that they would get the same results? (= the study replicates). List some actions the researcher took to enhance reliability.

Figure 3

9. Pilot: Was a pilot done? (2 points)
   10. Sample: (6 points)
      What is the sample size?
      What do you think of that sample size? (Support your answer)
      How was the sample selected? (random or non-random?)
      Is the sample representative of the population? (Support your answer)

11. Data collection: (6 points)
    How were the data collected? (Were surveys used? Was it self-report or were subjects observed? etc.)
    What were the circumstances of data collection?
    How many researchers were there?

12. Limitations: (6 points)
    List and evaluate the controls used in the design and data collection.
    Comment on each of the following validity threats:
    a. History:
    b. Maturation:
    c. Mortality:
    d. Self-selection:
    e. Testing:
    f. Instrumentation:
Author Information

Sally Northam, PhD, RN
Professor, College of Nursing and Health Sciences The University of Texas at Tyler 3900 University Blvd. Tyler, Texas

Janet Lakomy, PhD, RN
Adjunct Associate Professor, College of Nursing and Health Sciences The University of Texas at Tyler 3900 University Blvd. Tyler, Texas