Fort Hays State University

FHSU Scholars Repository

Summer 2010

ADN program benchmarking using standardized exams for assessment and remediation

Amy J. Waters
Fort Hays State University, ajwaters@scatcat.fhsu.edu

Follow this and additional works at: https://scholars.fhsu.edu/theses

Part of the Nursing Commons

ADN PROGRAM BENCHMARKING USING STANDARDIZED EXAMS

FOR ASSESSMENT AND REMEDIATION

being

A Thesis Presented to the Graduate Faculty

of the Fort Hays State University in

Partial Fulfillment of the Requirements for

the Degree of Master of Science in Nursing

by

Amy J. Waters

B.S.N., The University of Kansas

Date _____   Approved_____
                                                                Major Professor

                                                          Approved_____
                                                                Chair, Graduate Council

GRADUATE COMMITTEE APPROVAL

The Graduate Committee of Amy Waters

hereby approves her thesis as meeting partial

fulfillment of the requirements for the Degree of

Master of Science in Nursing

Date _____   Approved _____

Chair, Graduate Committee

Approved _____

Committee Member

Approved _____

Committee Member

Approved _____

Committee Member

Copyright © 2010 Amy J. Waters

ADN PROGRAM BENCHMARKING USING STANDARDIZED EXAMS FOR

ASSESSMENT AND REMEDIATION

Amy J. Waters, MSN(c)

Fort Hays State University, 2010

Supervising Professor: Dr. Karolyn Kells

ABSTRACT

The purpose of this research investigation was to determine the correlational values between testing scores when utilizing the Assessment Technologies Institute™, LLC (ATI) standardized content exams during the third and fourth semesters of an associate degree nursing (ADN) program, and the sub-scale scores in the same content areas of the comprehensive final exam, the ATI RN-Predictor, taken during the final weeks of nursing school. The RN-Predictor is used to assess graduating students' preparedness for the National Council Licensure Examination for Registered Nurses (NCLEX-RN®).

A Level II, nonexperimental, correlational study design utilized convenience and retrospective sampling of academic variable data of nursing program graduates ($N = 91$) of a Midwestern ADN program, including two campuses, from three academic years. Pearson correlation coefficients and Spearman rho were used to explore associations between the content scores and comprehensive sub-scale scores, and between the ATI RN Predictor scores and NCLEX-RN® outcome (dependent variable).

Research question number one was, "Are there statistically significant relationships between student ATI™ Content Mastery Series scores and ATI™ RN-Predictor sub-scale scores in the like content areas?" This research question included

the Content Mastery scores, taken after completion of didactic content during curriculum and RN-Predictor scores in the following areas: medical-surgical, nursing care of children, and mental health. A weak, direct significant correlation was found between medical-surgical assessment scores. A moderate, direct significant correlation was found between assessment scores in nursing care of children.

Research question number two was, "Is there a statistically significant relationship between the student ATI RN-Predictor Comprehensive score and NCLEX-RN® outcome?" A weak, direct significant correlation was identified between these variables.

Research question number three was, "Are there statistically significant relationships between student ATI Content Mastery Series™ scores and NCLEX-RN® outcome?" A moderate, direct significant correlation was identified between Content Mastery medical-surgical scores and NCLEX-RN outcome®. A weak, direct significant relationship was found between nursing care of children scores and NCLEX-RN outcome®. No significant correlation was found between mental health scores and NCLEX-RN outcome®. There were insufficient reported maternal newborn scores to analyze this correlation.

An additional ad hoc research question was: "Is there a statistically significantly relationship between ATI RN-Predictor Comprehensive sub-scale scores and NCLEX-RN® outcome?" No significant relationship was found between either medical-surgical scores or maternal newborn scores and NCLEX-RN® outcome. A weak, direct significant correlation was found between nursing care of children scores and NCLEX-

iv

RN® outcome. An inverse but not significant relationship was found between mental health scores and NCLEX-RN® outcome.

Findings from this investigation may be useful for establishing benchmark goals for progression and completion policies at this program. Results may also be beneficial for refining nursing program curricula at this site and facilitating identification of at-risk student testing areas leading to remediation before program completion and NCLEX-RN® testing.

ACKNOWLEDGMENTS

I wish to acknowledge my thesis chair, Dr. Karolyn Kells, for her continuous support and advice for thesis completion. Great gratitude also goes to my thesis committee members for their help, knowledge, and time during this investigation: Dr. Robert Meier, Dr. Carol Moore, and Professor Dianna Koerner.

Much thanks to research assistants at the investigation site, Joyce Jones and Karen Harden. This project would not have been possible without all of your hard work and hours spent gathering the data. Thank you!

Finally, thank you to the faculty and administration at Pratt Community College for your support and assistance to allow this research. I hope you benefit from the findings.

DEDICATION

This thesis is dedicated to my family and friends for the years of support and encouragement. To my parents, Bethel and the late Lavern Goltl, thanks so much for the financial and emotional support to enable this journey for my MSN. To my husband Mike, many thanks for your continuous support and love. To my children, Ellen, Owen and Brynn, thank you for your patience and encouragement through the semesters of thesis work.   I hope you will benefit from my pursuit and realize the importance of knowledge and discovery. To all of my dedicated friends and family not mentioned here, I appreciate your words of inspiration.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDIXES

CHAPTER 1 - INTRODUCTION

Nationally, nursing programs are continuously monitoring and adapting curricula to adhere to changing technology and advancing complexity of health issues. This never-ending task is needed to graduate students who will become competent registered nurses (RN). Passage of the National Council Licensure Examination for Registered Nurses (NCLEX-RN®) (National Council of State Boards of Nursing, 2007) is considered a measure of minimum standard level of competency for nursing practice at an entry level for licensure (Yin & Burger, 2003). In the United States, a program's pass rate for graduates on the NCLEX-RN® is considered a critical measurement of not only appropriate, but minimally adequate, nursing curriculum.

Statement of the Problem

Graduate nurse failure on the NCLEX-RN® can result in financial and emotional hardships (Briscoe & Anema, 1999; Yin & Burger, 2003). Additionally, research indicates that those graduates who fail the NCLEX-RN® experience decreasing probability of successful testing on repeated attempts (Zuzelo, 1999). Failure rates also reflect on nursing schools, which can resulting in decreased funding and loss of state approval for the program by the state nursing governing body (Kansas State Board of Nursing [KSBN], 2008) when a state imposed passage rate goal is not met (Morrison, Free, & Newman, 2002). These nursing schools also face scrutiny from their own governing boards when passage rates fall below state or national standards, limiting financial and advisory support. Poor passage rates affect the reputation of nursing programs and can impede accreditation. Success of graduate nurses on the NCLEX-RN®

1

is often viewed as a measure of program worth and quality (Beeman & Waterhouse, 2001; Daley, Kirkpatrick, Frazier, Chung, & Moser, 2003). Nationwide, NCLEX-RN® failure further contributes to the nursing shortage when graduating nurses cannot obtain licensure by successfully completing the NCLEX-RN® (Haas, Nugent, & Rule, 2004; Spurlock & Hanks, 2004; Wong & Wong, 1999). Nursing programs are guided in curriculum development by reviewing the most recent detailed test plans provided by the National Council of State Boards of Nursing (NCSBN) (2007), by trending curriculum changes and effects on students, by setting and measuring benchmarks, by expert consultant advice, and by trial and error. Literature lacks agreement on what nursing programs can implement to benchmark student progress toward first-attempt success on the NCLEX-RN® by graduates. The only agreement is that every school must form its own benchmarking process (Davenport, 2007; Fowles, 1992; Newton, Smith, Moore, & Magnan, 2007).

<div align="center">Purpose of the Investigation</div>

The purpose of this research investigation was to determine the correlational values between testing scores when utilizing the Assessment Technologies Institute™, (ATI) standardized content exams during the third and fourth semesters of an associate degree nursing (ADN) program, and the sub-scale scores in the same content areas of the comprehensive final exam, the ATI RN-Predictor (2007c), taken during the final weeks of nursing school. The RN-Predictor is used to assess graduating students' preparedness for the NCLEX-RN®.

Significance of the Investigation

Results from this investigation will be useful for nursing faculty in curriculum development, progression of students, and remediation for success. Identified relationships between academic variables of testing scores and predicted successful NCLEX-RN® testing can guide faculty to develop and refine student progression (Briscoe & Anema, 1999).

Nursing faculty can implement curriculum changes based on the findings to optimize academic achievement and improve NCLEX-RN® passage rates, thereby educating graduate nurses (GN) who are academically prepared to enter the work force. Before graduation at-risk nursing students can be identified for remediation, tutoring, and counseling interventions to improve in content areas below mastery, thus becoming better prepared to NCLEX-RN® test (Arathuzik & Aber, 1998; Briscoe & Anema, 1999; Haas et al., 2004; Siktberg & Dillard, 2001). Results gained from this investigation will create a baseline for the ADN program investigation setting to be utilized for continued tracking and trending of data for benchmarking.

Theoretical Framework

Imogene King's (1971) Interacting Systems Framework and Theory in Nursing Practice concepts are organized around systems, from personal and interpersonal to larger social systems, such as community and educational institutions. King's framework expanded to incorporate the Theory of Goal Attainment (1981) and the concept of learning. Ongoing clarification and theoretical discussion of King's interacting systems framework over the years have included contemporary topics of technological advances

and changes in organization. This would be applicable to curriculum development and standardized computerized testing, such as the NCLEX-RN®, and for research and remediation for successful goal attainment.

Foundational to her Personal Systems Framework, the concept of growth and development is defined by King (1981) as "the processes that take place in an individual's life that help the individual move from potential capacity for achievement to self-actualization" (p. 31). King's (1986) concept of learning is defined as "a process of perception, conceptualization, and critical thinking involving multiple experiences in which changes in concepts, skills, symbols, habits, and values can be evaluated in observable behaviors and inferred from behavioral manifestations" (p. 24). King (1981) defined the concept of organization within her Social Systems framework as "a system whose continuous activities are conducted to achieve goals" (p. 119). King identified experiences of concepts, skills, critical thinking and values in her theory of learning and they are evident in the teaching-learning process. For individuals who set the goal of becoming nurses, King would consider the personal growth and development of moving from potential to reality through achievement as part of her Personal Systems framework and Theory of Goal Attainment. The process of continuous improvement in nursing knowledge and education are inherent in the concept of organization in the Social Systems framework. This continuous improvement involves tracking and trending data for the short-term use of identification of at-risk areas and students and for the long-term use of curriculum development. King's Theory of Goal Attainment applies to the larger

social systems of nursing programs using standardized assessment testing for benchmarking. Figure 1 displays King's foundational impact on the variables of interest.

<center>Definitions of Variables</center>

The variables of interest for this investigation were selected based on a literature review and faculty discussion of these variables regarding curriculum development, progression and remediation of nursing students. Both dependent and independent variables will be discussed.

ORGANIZATION

LEARNING

ATI CONTENT MASTERY SCORES

REMEDIATION

ATI RN-PREDICTOR SCORES

GROWTH & DEVELOPMENT

RN

NCLEX-RN

PERSONAL

SOCIAL

GOAL
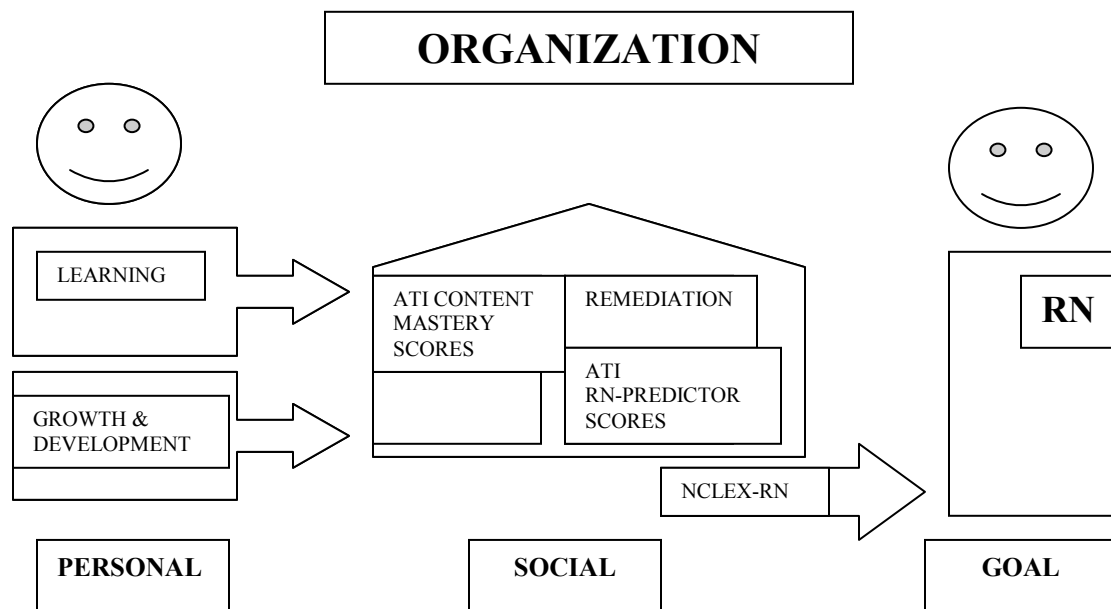
Figure 1. Schematic Model of Investigational Framework

Dependent Variable

The dependent variable for this investigation was the NCLEX-RN® (2007) outcome. Theoretically, the NCLEX-RN® is the national licensing exam that all graduate nurses must complete successfully to be eligible for licensure as an RN.

Operationally, the NCLEX-RN® outcome is pass or fail. This data was collected from nursing program records.

Independent Variables

Although the potential independent variables were almost limitless, the research setting selected student scores that were useful for this particular investigation.

Assessment Technology Institute Diagnostic Assessment Scores

Theoretically, the Assessment Technology Institute (ATI) Content Mastery Series 2.1 diagnostic assessments are classified according to the current NCLEX-RN® test plan (Assessment Technologies Institute™ [ATI], 2005). The ATI updated the Content Mastery series assessments to align with NCLEX-RN changes during the time span of this research. Content Mastery 2007 assessments were administered to the sample group starting in the 2008-2009 academic year.

Operationally, the scores were provided in composite scores for individuals and group averages, based on cognitive levels, outcomes, and NCLEX-RN® (2007) categories. The ATI also provides classifications of Criterion-Referenced Proficiency Levels for individual scores, developed based on a national standard setting study (ATI, 2007). The assessment test was administered to students in a proctored setting following regular didactic instruction to establish their mastery of content. For the purposes of this investigation, the following areas of content were examined for correlation with the ATI Comprehensive RN-Predictor sub-scale student scores: Maternal-Newborn Nursing Care (MN), Nursing Care of Children (NC), Mental Health Nursing (MHN), and Medical-Surgical Nursing (MS).

ATI Comprehensive RN-Predictor

Theoretically, the Comprehensive RN-Predictor 3.0 Assessment sub-scale scores assess students' preparedness for the NCLEX-RN® (ATI, 2005). The RN-Predictor test was updated by the company during the research time frame to align with the NCLEX-RN® standards. The 2007 RN-Predictor was administered at this site starting in 2008.

Operationally, the ATI RN-Predictor scores were provided in composite individual scores. The assessment contained 180-items designed in the NCLEX-RN® format, utilizing alternate-item testing. The comprehensive assessment provided a student with areas of needed remediation based on incorrect assessment responses by subtopics. The total composite score for each student were used for this variable. Sub-scale scores in the content areas of interest were provided as composite scores for individuals and group averages. These composite individual scores were used for comparison to the ATI Content Mastery Series 2.1 diagnostic assessments scores in like content areas.

Research Questions

For this investigation, the research questions examined were as follows:

1. Are there statistically significant relationships between student ATI Content Mastery Series™ scores and ATI RN-Predictor sub-scale scores in the like content areas?

2. Is there a statistically significant relationship between the student ATI RN-Predictor Comprehensive score and NCLEX-RN® outcome?

3. Are there statistically significant relationships between student ATI Content

   Mastery Series™ scores and NCLEX-RN® outcome?

Assumptions

Assumptions for this investigation included the following:

1. It was assumed that data collection of the academic institution was complete

   and accurate.

2. It was assumed that the program didactic content for content areas of student

   examination preceded ATI assessment testing.

3. It was assumed that students have put forth effort into achieving accurate

   assessment scores on all ATI examinations.

4. It was assumed that students did not utilize or had access to ATI Online

   Practice Assessment until initial mastery testing has occurred.

Delimitations

The investigation delimited the parameters of this investigation to the following:

1. Online LPN to RN nursing students of the program were delimited for

   statistical equivalence reasons due to their limited use of proctored ATI

   Content Mastery Series™ testing.

2. All students that repeated a semester in the nursing program were delimited

   from the investigation since their scores had potential to skew the data.

Limitations

Limitations for this investigation showed that the results cannot be generalized to

all nursing programs due to convenience and retrospective sampling. The investigation

setting used was in a rural state geographic location, so results cannot be generalized to urban schools. The sample nursing program was a small rural ADN college; therefore, the findings of the investigation cannot be generalized to baccalaureate or diploma nursing students.

Summary

Failure on the NCLEX-RN® examination for licensure after successful nursing program completion is a national issue that affects individuals, nursing schools, state boards of nursing, and health care facilities. The growing populations of aging citizens and those with complex health problems are already faced with a dramatic shortage of nurses. The number of GN's that are not successful on NCLEX-RN® make this situation even more drastic. Nursing programs need to assure that they are addressing the needs of healthcare nationwide and the needs of their own nursing students by continuous review of curricula. Use of assessment testing and remediation plans led by benchmarking goals for students are steps that programs can explore for increasing student success on the NCLEX-RN®. Investigations such as this one can help nursing schools identify and address those needs that will lead to NCLEX-RN® success.

CHAPTER II-REVIEW OF LITERATURE

Chapter I addressed the need for research exploring the use of standardized assessment testing for benchmarking in nursing programs. A Level II quantitative, non-experimental, correlational, ex-post facto investigation was proposed. Chapter II will examine the nursing literature base concerning use of standardized assessment testing to set program benchmark goals for success and to identify nursing students at risk for failure on the National Council of Licensure for Registered Nurses exam (NCLEX-RN®). An introduction to nursing program benchmarking and a summary of standardized testing used for benchmarking for student program and NCLEX-RN® success are discussed.

Student failure on the NCLEX-RN® examination affects individuals emotionally and financially. Financial burden accompanies individuals who are not successful on first-attempt NCLEX-RN® testing in the form of unemployment, remediation fees and retesting fees. Institutionally, nursing programs suffer loss of funding, accreditation, and reputation when licensure passage rate goals are not met (Giddens, 2009; Morrison et al., 2002). The nationwide nursing shortage worsens with NCLEX-RN® failures, impacting staffing and safety in health care facilities.

Nursing programs continually strive to find the evidence supporting curriculum changes to meet the needs of student nurses for success. Graduating from an accredited nursing program with high grades does not guarantee first-attempt success on the NCLEX-RN®. Vast research has been conducted throughout the years to predict which modifiable factors result in successful program completion and first-time NCLEX-RN® success. Predictors of interest have included student demographics and personality

factors, as well as academic variables, such as prerequisite courses and grades, grade point averages, course credit loads, and assessment scores (Campbell & Dickson, 1996; Mosser, Williams, & Wood, 2006; Waterhouse & Beeman, 2003). Little research has been conclusive for providing a basis for programs to change. Each nursing program is responsible for tracking student data to identify areas for improvement (Fowles, 1992).

Benchmarking

Benchmarks are the measure of a best practice; benchmarking is the process of identifying benchmarks and applying them for performance improvement (Billings, 2007). Benchmarking is reported to have begun by corporations in the early 1980's to improve performance and drive competition (Anonymous, 2008). A reference to benchmarking as early as 1989 described the process as "ongoing" and "a systematic approach" (Spann, 1997).

Benchmarking started in the health care industry to reform and improve performance of health care facilities to meet quality improvement standards. The process of setting benchmarks leads to a search for best performers or practices for optimal patient outcomes (Spann, 1997). Improving quality of care has been a vital component in the nursing profession since Florence Nightingale's work.

In nursing education, benchmarking has been implemented to identify best practices of teaching and learning toward success and competence. Continuous improvement of curriculum and student outcomes requires a systematic approach to determine best practice. "Benchmarks assist in evaluating the extent to which schools of nursing promote and sustain excellence over time" (Scheckel & Valiga, 2007, p. 167).

Continuous quality improvement and educational research are two of the 30 National League for Nursing (NLN) Hallmarks of Excellence in Nursing Education© (National League for Nursing, 2009). These Hallmarks are considered characteristics or traits for nursing faculty to demonstrate professionalism toward outstanding achievement.

The process of benchmarking is an effective strategy to address a variety of nursing education concerns. Benchmarked data are useful for improving program outcomes and facilitating the accreditation process (Billings, 2007). Benchmarking can also be utilized at course levels to evaluate content areas and at individual levels by faculty for ongoing self-evaluation of instruction.

Research of Standardized Exams

Fowles (1992) conducted a retrospective study to identify predictors of NCLEX-RN success and program completion at an upper-division, baccalaureate nursing program. Convenience sampling of graduated student data from 1985 to 1988 was used for a total of 192 students. The NCLEX-RN assessment format for this study was paper and pencil format, not computerized.

Numerous preadmission independent variables were examined by Fowles (1992) for correlation, including American College Testing Assessment (ACT) scores: composite (ACTCOMP), English (ACTENG), math (ACTMATH), social science (ACTSOCSC), and natural science (ACTNTSCI); lower division grade point average (GPA) after prerequisite completion (GPAPRERQ), lower-division GPA in science courses (SCIENGPA), lower-division GPA in liberal arts courses (LBARTGPA), grade in Anatomy and Physiology I (GRDANP1), grade in Anatomy and Physiology II

(GRDANP2), GPA at end of nursing curriculum Level I (GPALEV1); GPA at end of nursing curriculum Level 2 (GPALEV2); and Mosby Assesstest percentage correct and percentile (PERCENT, PCTILE). Dependent variables were NCLEX-RN (NCLEX) scores and cumulative GPA at end of nursing curriculum (GPALEV3).

Instruments used for the Fowles (1992) study were student data, the ACT, the Mosby Assesstest, and the NCLEX-RN. Stepwise multiple regression analyses were conducted to generate a combination of variables that best reflected success in the program and on the NCLEX-RN. Analyses were run to control for interdependence between ACT subscale and composite scores and for the effects of cumulative GPAs within the nursing program. Level of significance was set at $p = 0.05$. Scores were available for 182 of the NCLEX results; five were reported only as pass/fail. Raw scores were converted to pass = 1600 and fail < 1599 status for operations. Findings showed all correlations were stronger when raw scores were used rather than pass or fail data.

Findings by Fowles (1992) of predictors for program success revealed strong interdependence between ACTCOMP, ACT subscales, and GPA within nursing. After regression analysis, GPALEV1, PCTILE, and GRDANP2 were found to be significant. Findings by Fowles (1992) for correlations of NCLEX-RN success revealed predictive strength for GPALEV1, PCTILE, and ACTSOCSC.

A strengths of this study was the use of appropriate statistics for correlating data variables. The study described maintenance of confidentially for student data collection. A weakness included limited generalizability to other program types and locations due to single site sampling. Instrument reliability and validity were not addressed.

Though felt to be essential for accounting for variance in Fowles (1992) findings, raw scores on the NCLEX-RN are no longer available for predictive effectiveness. This study has relevance to the proposed investigation for its main purpose of determining predictive factors in relation to nursing student success within nursing program and on the NCLEX-RN. One variable examined was a standardized computerized exam, the Mosby Assesstest. The Mosby Assesstest percentile scores were found to be significant for both student program and NCLEX-RN success.

Alexander and Brophy (1997) conducted a retrospective, correlational study over five years at a baccalaureate nursing program to identify admission, progression and exit variables that were predictive of NCLEX-RN results. A quota sampling technique of convenience was used to gather data from July 1988 through February 1994 ($N = 188$). Student data included all NCLEX-RN first attempt failures ($n = 94$) for the time frame and 94 randomly selected graduates who experienced NCLEX-RN first attempt success ($n = 94$).

The dependent variable for Alexander and Brophy's (1997) study was NCLEX-RN performance. Independent variables included admission variables of high school rank, SAT scores, years of high school chemistry and math, admission status, GPA, number of credits, and age. Independent variables of progression were nursing course grades, nine cognate courses, and Level 1 and Level 2 GPA within the nursing curriculum. Nursing courses included Nursing Fundamentals, Childbearing Nursing Adult I, Child/Family, Nursing Adult II, Mental Health, and Nursing Seminar. Cognate courses for this study included Introductory Psychology, Anatomy and Physiology I

and II, Communication, Developmental Psychology, Microbiology, Elementary Chemistry, Organic Chemistry, and Introductory Sociology. The independent exit variable of study was the National League for Nursing (NLN) Comprehensive Achievement Test student score, taken at the end of year 2. Instruments used were student records and the NLN Comprehensive Achievement Test.

Alexander and Brophy (1997) performed descriptive and inferential analyses. To determine mean differences between groups on each of the independent variables, t-tests were computed. By grouping nursing and cognate variables, t-tests were used to compare between the groups. Chi-square analysis was used to test for proportionality of course failures between the groups. Logistic regression analyses were performed to indicate the strongest predictors of NCLEX-RN success.

Statistically significant differences were revealed for course grades between those who passed and those who failed NCLEX-RN. The greatest magnitude of difference between the two groups occurred with the nursing courses and GPA at Levels 1 and 2. Grades in five of the six clinical nursing courses provided the greatest differentiation between the two groups of pass and fail; mental health clinical grade was the exception. Results of the t-tests comparing the groups on the progression variables remained statistically significant; Nursing Seminar was the sole exception. There were statistically significant differences found between the two groups of NCLEX-RN outcomes on the NLN test, which is consistent with findings from previous research.

Alexander and Brophy (1997) found significant differences in proportions of courses passed and failed by the successful and unsuccessful NCLEX-RN graduates,

$X^2 = 44.61$, $p = .001$. Of the 94 graduates who failed NCLEX-RN, 19 failed nursing and cognate courses, 26 failed nursing courses, and 18 failed cognate courses; 31 of the fail group had no course failures. Three logistic regression stepwise models were tested; the finished model correctly predicted 88.24% and used only two courses for prediction: Nursing Adult I and Introductory Sociology. The second model tested the fit of SAT verbal and SAT math scores as predictors: the ability to predict accurately was 68.22% with only SAT verbal remaining in the final model. The third model of logistic regression analysis tested the fit of six nursing course grades and the NLN test score. This model included the NLN test in combination with three nursing courses: Childbearing, Nursing Adult I, and Mental Health. Accurate predictive ability was 80.63%.

Strengths of Alexander and Brophy's (1997) include appropriate statistical analysis used for the levels of measurement for the variables. A gap in research was identified for this time period: the revised NCLEX-RN test plan was implemented in 1988 at the beginning of this study. Variables used for this study were varied, identifying three points in the nursing program to explore for predictors. Weaknesses include that generalizability is limited due to single site research. Instrument reliability and validity information was not provided. Also, there was a lack of information regarding the data collection process and maintenance of confidentiality.

This study (Alexander & Brophy, 1997) was useful to the investigation based on the findings of predictors for NCLEX-RN success. Although the variables used for this study are not identical to the proposed variables of interest, findings identified significant

predictors at key points within the nursing curriculum: admission, progression and exit. One standardized exit exam, the NLN Comprehensive Achievement Examination, was found significantly predictive for graduate NCLEX-RN success per logistic regression analysis.

Briscoe and Anema (1999) conducted a retrospective, correlational study at an associate degree nursing (ADN) program in a public urban university located in the southern United States to examine selected academic and non-academic variables for prediction on NCLEX-RN outcomes. Convenience sampling of existing student records of May, 1997 ADN graduates were used for total sample size of 38. Data were collected utilizing an instrument developed by the investigators, the Student Profile Form. Six dependent variables were correlated with the independent variable of NCLEX-RN outcomes: pre-admission GPA, grades in clinical nursing courses, scores on the National League for Nursing (NLN) Comprehensive Achievement Test I and II, age, and race. The NLN tests I and II and NCLEX-RN exam were also instruments used in Briscoe and Anema's research.

Briscoe and Anema (1999) used six hypotheses for the study. Hypothesis one addressed the difference between students with pre-admission GPAs of 2.25 – 2.45 to students with pre-admission GPAs of > 2.5. There was no significant difference found between these student groups, thus the hypothesis was rejected. Hypothesis two addressed the relationship between failing a nursing course and NCLEX-RN outcome. No significant relationship was found; therefore, the hypothesis was rejected. Hypothesis three addressed the relationship between scores on the NLN I exam and NCLEX-RN

outcome: findings revealed a significant relationship at the .01 level, $r = .476$. This hypothesis was accepted as valid. Hypothesis four addressed the relationship between scores on the NLN II exam and NCLEX-RN outcome. There was a significant relationship at the .01 level, $r = .371$; this hypothesis was accepted. Hypothesis five addressed the relationship between graduate age and NCLEX-RN outcome: findings revealed a significance at the .05 level, $r = .373$. This hypothesis was accepted. Hypothesis six addressed the relationship between race and NCLEX-RN outcome. Race was categorized into four groups: White non-Hispanic, Black non-Hispanic, African descent, and Hispanic. There was a significant relationship at the .05 level for African descent students, $r = .471$; thus the hypothesis for that group was accepted. No significant correlations were found for the other race categories.

Strengths for Briscoe and Anema's (1999) study include appropriate statistical analysis for the level of measurement for the variables. The study design was appropriate for the intent of correlating student data. Weaknesses include the inclusion of a name column on the Student Profile Form which inhibits protection of confidentiality. No measures for maintenance of confidentiality were addressed in the data collection methodology. Instrument reliability and validity measures were not provided. Generalizability is limited due to single site research and small sample size of only ADN graduates.

Findings of this study (Briscoe & Anema, 1999) support previous research findings revealing the predictive effects of the NLN Comprehensive Assessment exams (Alexander & Brophy, 1997; Campbell & Dickson, 1996). The use of standardized

exams for assessment and remediation has been identified as a useful strategy. With early identification of NCLEX-RN outcomes prediction, timely and meaningful interventions can target students at-risk for failure.

Washington and Perkel (2001) conducted a retrospective pilot study to investigate factors contributing to NCLEX-RN outcome at a university-based nursing program. Convenience sampling was used for data collection of students completing graduation requirements in May 1998 ($N = 67$). The sample was divided according to program types: basic option graduates ($n = 47$), and accelerated option graduates ($n = 20$). Instruments used for this study were student records and the Arnett Pre-RN Readiness Examination (Arnett). The NCLEX-RN outcome was the dependent variable.

For Washington and Perkel's (2001) study, variables of ethnicity, language, age, and gender were cross-tabulated with transfer GPA, cumulative GPA, repeated science courses, repeated nursing courses, Arnett test scores, and NCLEX-RN results for both groups. Among the basic option graduates, ethnicity was significantly correlated with cumulative GPA, $p = .012$, and repeated science courses, $p = .002$. For the accelerated student sample, significant correlations were revealed between age and transfer GPA, $p = .047$, and between language and transfer GPA, $p = .049$. For both groups, an overall mean transfer GPA was 2.96 and a mean cumulative GPA of 3.2.

Washington and Perkel (2001) performed a correlation analysis among transfer and cumulative GPA, repeated nursing and science course, Arnett scores and NCLEX-RN results in each group using one-tailed and two-tailed directionality. At the .05 level of significance, and using one-tailed test with the accelerated option sample,

significant correlations occurred between transfer GPA and Arnett test, $p = .027$, and between cumulative GPA and Arnett test, $p = .010$. In the basic option sample, at the .05 level of significance, transfer GPA and repeated sciences were significantly correlated, $p = .014$; and cumulative GPA and repeated sciences, $p = .012$. At the .01 level of significance, the basic options group showed significant correlations between repeated nursing and repeated sciences, $p = .007$.

Washington and Perkel (2001) performed an analysis of variance (ANOVA) to explore the differences between outcome on the NCLEX-RN and variables. Significant differences at the .05 level were found between transfer GPA and Arnett test scores, $p = .027$, and between cumulative GPA and Arnett test scores, $p = .010$. Using logistical regression analysis of forward stepwise regression, the Arnett test was the only significant variable found to predict NCLEX-RN results, $p = .012$. Using ANOVA, there were no significant difference between NCLEX-RN outcomes and the other variables.

Strengths of Washington and Perkel's (2001) study include the appropriate statistical analysis for the variables and levels of measurement. The study proposal process for this study was adequately described. Weaknesses of this study include that generalizability is limited due to small sample size and single site research. Instrument reliability and validity for the Arnett was not addressed.

Findings from Washington and Perkel's (2001) study revealed a significant prediction between Arnett testing and NCLEX-RN. Though no variables of this study are variables of interest for the current investigation, it indicates the importance of identifying predictive variables for remediation measures.

Uyehara, Magnussen, Itano, and Zhang (2007) examined potential predictors of program and NCLEX-RN success, and program withdrawal in a study to evaluate program changes addressing attrition and NCLEX-RN success rates. Student data at a generic baccalaureate program in Hawaii were collected over a five-year period following a newly implemented curriculum. Three phases for significant data collection were selected for this six semester baccalaureate program: admission data, within program data, and exit predictor data. Dependent variables for this study were program success and withdrawal and NCLEX-RN outcome. Program success was considered completing the program within one semester of expected completion, while program withdrawal was considered leaving the program before graduation.

Independent admission data for this study (Uyehara et al., 2007) included prerequisite grade point average (GPA), cumulative GPA, National League for Nursing (NLN) pre-nursing verbal, math, science, and composite scores, and ethnicity. Independent data variables during the nursing program included grades from the Pathophysiology, Fundamentals, Adult Health Nursing I, Mental Health Nursing, Maternal-Newborn Nursing, Pediatric Nursing, and Adult Health Nursing II courses, the NLN achievement test scores for mental health, childbearing family, pediatric nursing, and adult health nursing, and the Watson Glaser Critical Thinking Appraisal. Independent exit predictor variables for analysis included Mosby Assess Test score, and the student nursing GPA at graduation. Instruments used were student records, the NLN tests, the Mosby Assess Test, the Watson Glaser Critical Thinking Appraisal, and the NCLEX-RN.

For this study (Uyehara et al., 2007), ten groups of admitted students in the five-year period of study were tracked until program completion or withdrawal. Convenience sampling of 280 generic baccalaureate students was used, resulting in a total 224 students with program completion. Data analysis included descriptive statistics, correlations, and regression analysis to determine significant relationships between the variables.

Uyehara et al.(2007) found that for NCLEX-RN outcome, significant correlations were found with the Mosby Assess Test ($n = 216$), $r = .24$, $p = .0003$, the NLN Adult Health Comprehensive Test ($n = 217$), $r = .41$, $p < .0001$, the NLN Maternal-Newborn Comprehensive Test ($n = 217$), $r = .16$, $p = .0179$, the NLN Pediatric Nursing Comprehensive Test ($n = 216$), $r = .20$, $p = .0025$, nursing GPA ($n = 217$), $r = .186$, $p = .0059$, and course grades in the Fundamentals course ($n = 217$), $r = .195$, $p = .0038$. Upon logistic regression analysis, the only significant predictor found was the NLN Adult Health Comprehensive Test, $p = .0001$. For program success, 213 students met the definition. There were no significant correlations among dependent and independent variables and no significant predictors of program success and withdrawal and NCLEX-RN outcome. Fifty-six students met the definition of program withdrawal. The only significant correlation for independent variables was course grades for pathophysiology ($n = 217$), $p = .0001$; the higher the course grade, the higher the probability of completion. Twenty-seven (48.21 %) of the 56 withdrawn students had a grade of C or below in the Pathophysiology course.

Strengths of this study (Uyehara et al., 2007) were the inclusion of independent variables collected at various points of program matriculation. Earlier identification of

students at risk for program withdrawal and failure can lead to earlier management. This program implemented a strategy policy of notifying those at-risk and high-risk students based on nursing course grades, standardized achievement test scores and nursing GPA. Other at-risk identifiers implemented by this program included the use of pathophysiology course grades and NLN Adult Health Comprehensive Test scores. An identified weakness is that generalizability is limited due to single site research. Information from this study is beneficial to the proposed investigation for leading to further exploration of using standardized assessment testing for setting program benchmarks and early identification of at-risk students.

<div align="center">Research of HESI Exam(s)</div>

Morrison et al. (2002) conducted a qualitative study to explore use of remediation strategies following adoption of the Health Education Systems, Incorporated (HESI) Exit Exam ™ at five nursing programs. The instrument used for   this study was the E^2, consisting of 160 test items. The HESI Exit Exam (E^2) was reported as a computerized, comprehensive exam that uses a predictive mathematical model to compare individual students with other E^2 participants across the United States. Based on the test scores, learning needs could be identified to design individualized remediation plans.

Morrison et al. (2002) used a convenience sample consisting of seven programs at five schools of nursing that had implemented a progression and remediation policy based on E^2 scores. Administrators at these programs were interviewed regarding NCLEX-RN

after using the remediation policy compared to pre-policy. Other areas of the interviews related to use of remediation as outlined in the various policies.

Morrison et al. (2002) findings for this study indicated that the implemented progression and remediation policies were effective for these programs: NCLEX-RN pass rates improved in all programs studied by 9-41% within two years after implementation of the policies. To further examine this increase in pass rates, the researchers applied a quantitative examination using a chi square test of significance to analyze pass rate per program before and after policy implementation. The chi square revealed significant increases in six of the seven programs, five programs at the $p = .001$ level of significance; one program at the $p = .05$ level of significance. A t-test for related samples was computed to examine all seven programs; NCLEX-RN pass rates were found to be significantly higher, $p = .002$ after implementation of progression policies.

Administrators at the programs were interviewed regarding remediation practices for the Morrison et al. (2002) study. According to summarized interview reports, remediation strategies varied per program. Recommendations for remediation at one program included instructing the students to study using the E^2 score printouts to identify subject areas independently. Another program required students to meet with the associate dean for support and advice. One program had students meet as a group with faculty to assist in collaborative review and computer assisted instruction. In one program, faculty re-taught a content area that was revealed to be weak throughout the student population. All administrators reported that students often failed to use remediative resources available.

Strengths of this study (Morrison et al., 2002) included the use of multiple settings, and different nursing program types. A gap in nursing literature was identified: there was not previous research with significant findings examining different program types in relation to the E². Weaknesses include poor transferability due to limited sample number. Despite the inclusiveness of different ADN ($n = 3$) and BSN ($n = 4$) programs, no diploma nursing programs were included in this study. Geographical information for the programs used was not provided to assess transferability due to regional differences. Trustworthiness for this study is difficult to assess as there was no discussion of audit trails, member-checking, number of researchers or teams for separate information checking.

Morrison et al. (2002) study was useful to the proposed investigation because it examined a standardized predictive exam as a variable for correlation with NCLEX-RN outcomes. It also attempted to explain how the HESI Exit Exam was being used as a benchmark for progression and remediation. The effectiveness of the remediation policies emphasizes the importance of identifying areas of learning needs for timely and guided improvement.

Daley et al. (2003) conducted an ex-post facto, correlational study to examine demographic and academic characteristics of students taking the NCLEX-RN in 1999 and 2000 to identify variables for prediction of NCLEX-RN test success. Two standardized tests, the Mosby AssessTest and the HESI Exit Examination, were independent variables studied for predictive ability in identification of students at-risk for failure on the NCLEX-RN. These assessments were both given to graduating

baccalaureate nursing students. Other independent variables examined were demographic variables of age, gender, and ethnic background, prerequisite grade point average (GPA), and American College Testing Assessment (ACT) scores. Pre-requisite program variables examined for correlation were prerequisite course grades in one organic and one inorganic chemistry course; a human anatomy course; an introduction to sociology course; and a zoology course providing human physiology. Nursing program variables identified for correlation in this study were: the nursing pathophysiology course, the senior medical-surgical nursing course, the clinical senior medical-surgical course, and the final cumulative program GPA. The dependent variable was the students' NCLEX-RN test result. Instruments utilized were the Mosby AssessTest, the HESI Exit Examination, the ACT, and the NCLEX-RN.

Daley et al. (2003) used convenience sampling, which included two classes of baccalaureate students: $n = 121$ in 1999 and $n = 103$ in 2000, for a total $N = 224$ for most variables. In 1999, the graduating senior group ($n = 121$) was required to take the Mosby AssessTest. In 2000, the graduating senior group ($n = 103$), was strongly encouraged to take the HESI Exit Examination ($n = 80$). Independent t-tests or chi-square tests were used to examine difference between variables. The significance level for this investigation was set a priori at $p < .05$.

Daley et al. (2003) reported that no significant differences were found between the two graduating groups, $p > .05$, for all variables. Overall, the HESI Exit Examination demonstrated greater test efficiency (91% HESI, 60 % Mosby); with greater predictive value compared to the Mosby AssessTest, both in negative (100% HESI, 97% Mosby)

and positive (22% HESI, 19% Mosby) prediction. Sensitivity (100% HESI, 85% Mosby)

and specificity (91% HESI, 57% Mosby) were also greater for the HESI Exit Exam group

(n = 80). In the 1999 graduating group, statistically significant findings between

NCLEX-RN passing and failing students included: older age ($p < .001$); higher

prerequisite GPA ($p < .005$); and higher ACT scores ($p < .005$). In the 2000 graduating

group, ethnicity was the only statistically significant variable: 33% of non-White

students were unsuccessful on the NCLEX-RN, compared to 4% of White students.

Validity on this finding is questioned due to small number of non-White students in the

sample. Program variables with statistically significant findings in the 1999 group were

anatomy grades, $p = .009$; nursing program pathophysiology grade, $p < .001$; didactic

nursing medical-surgical course grade, $p < .001$; clinical nursing medical-surgical course

grade, $p < .001$; and final cumulative GPA, $p < .001$. Two program variables for the

2000 graduating group were statistically significant: didactic nursing medical-surgical

course grade, $p = .004$, and final cumulative GPA, $p = .04$. Students who voluntarily

took the HESI Exit Examination in 2000 were statistically significant when correlated for

NCLEX-RN® results, $p = 0.01$.

A strength of the Daley et al. (2003) study is the comparison of standardized

tests in predictive nature for their purpose; the findings led to making the current and

more predictive HESI Exit Examination mandatory for graduates at this program. A

weakness addressed was the low sample number of non-White participants, making the

variable validity questionable. Another weakness identified is the unequal number of

students in the groups, and some incomplete data related to student age and ACT scores

when transferring from other schools. Instrument validity and reliability was not specifically addressed.

The Daley et al. (2003) study identified statistically significant findings of variables of concern for the sample, which confirms previous research. As this study illustrates, use of predictive program variables and efficient standardized testing scores allows faculty to identify vulnerable students and make curriculum changes.

Nibert, Young, and Britt (2003) examined new data from a previous study, with the intent of expanding the original report exploring use of the HESI Exit Examination (E^2), administered in the final semester of nursing programs. Goals of this ex-post facto, quasi-statistical, qualitative study were to describe progression policies by nursing programs that utilized the E^2 during the academic year 1999-2000; identify E^2 scores used as benchmarks for progression; and explore remediation strategies designed to complement the E^2 scores for benchmarking according to the progression policies. Seven-item questionnaires, which served as an instrument (Nibert, Young, & Adamson, 2002) were mailed to administrators of 166 schools of nursing that administered the E^2 to 6300 students. Data for this study were obtained from the original study survey by Nibert et al. Other instruments used included the E^2 student scores. Additional data were obtained from the HESI database regarding different versions of the E^2 to retest low-scoring students. Participating respondents for the questionnaire provided a sample ($N = 156$); 92 ADN programs, 63 baccalaureate programs, and three diploma nursing programs.

When reviewing findings from surveys regarding adoption of a progression policy, Nibert et al. (2003) found that 45 (30.20%) of $N = 156$ schools reported implementation or maintenance of a progression policy during the study year. Three consequences most commonly identified for students who achieved below the benchmark E^2 score were: denial or eligibility for graduation (51.43%), an incomplete or failing grade in the capstone course (34.29%), and/or withholding of approval for NCLEX-RN candidacy (14.29%). Additional data from HESI scores resulted in mandatory retesting for those students below identified benchmarks at all 45 programs with progression policies. In the study year, 36 (80%) of the schools used a different version of the E^2 for this purpose. Mandatory retesting policies were implemented at four additional programs the following academic year. Of the 35 progression policies submitted to the survey researchers for review, 20 (57.14%) specified the number of allowable retesting attempts; all only with the different version of the test. Seven schools permitted one retest, nine schools (45%) permitted two retests, and four school permitted unlimited retest opportunities; fees paid by students.

For benchmarking the E^2 scores, Nibert et al. (2003) found that only 45 (30.20%) of the schools reported that the progression policy designated a specific E^2 benchmark score for progression. Scores ranged from 77 to 90, with a majority (80%) using 85 as the benchmark score.

Concerning the progression policy and designated mandatory remediation for progression, Nibert et al. (2003) reported that 149 (94.30%) of the surveyed administrators responded. Most program administrators (71.81%) reported that

remediation was not required, but recommended. Strategies recommended for both mandatory and recommended remediation for students achieving below benchmark scores included: a specially designed course (52.38%); completion of computer-assisted instruction programs (23.81%); a comprehensive review guided by NCLEX preparation books (9.52%); participation in faculty tutoring sessions (9.52%); completion of an NCLEX simulation exam (2.38%); and re-enrollment in core nursing courses (2.38%).

Strengths for this study (Nibert et al., 2003) included new and gathered data that filled a gap of information for schools that utilize the HESI E². Consideration of the progression policies in consideration and adoption at other programs can guide faculty to formulate progression policies to fit their needs. Another strength was the inclusion of various programs types. A weakness identified is the impaired generalization based on differing student class sizes. Programs graduating few students can tolerate less risk of NCLEX-RN failure, and may want to choose a conservative approach in selecting a benchmark score (Nibert et al., 2003). Another major weakness is the inability to assess trustworthiness of this study due to the lack of details regarding data analysis. It was not reported how many teams of researchers for independent review participated; nor use of inquiry audit or member-checking of the data. The article does indicate multiple triangulation: person triangulation, involving data from multiple institutions using the HESI E², and investigator triangulation, although the number of trained researchers was not clearly stated. Use of data triangulation does increase the credibility factor for this study.

Areas of qualitative data collected by Nibert et al. (2003) are beneficial to nursing programs when formulating and adopting progression policies for students. Identifying set benchmarks and evaluation of outcomes of scores and progression policies leads to an enhanced ability to assist students in becoming candidates for the NCLEX-RN, as well as, improve program pass rates.

Higgins (2005) conducted a quantitative and qualitative study in three phases at a community college nursing program in Texas to determine strategies to raise the student NCLEX-RN® pass rate and lower the attrition rate. The study was conducted in three phases.

In Phase 1 of the Higgins (2005) study, ex-post facto data were collected from former nursing student records from fall 1999, spring 2000, and fall 2000 for a total sample of 213 students. Convenience sampling was used to include all enrolled students. Relationships between the dependent variables of program completion and/or NCLEX-RN® outcomes and the independent variables were explored. These variables included: prerequisite courses, preadmission test components, demographic variables, Health Education Systems, Incorporated (HESI) Exit Examination score, and nursing skills laboratory scores. Instruments used were the preadmission test, the HESI Exit Exam, the NCLEX-RN®, and the student data records. Correlation coefficients were run to test for significant relationships, with a chosen alpha level for rejection of .05; the region of rejection was two-tailed.

Regarding prerequisite course grades, Higgins (2005) found statistically significant differences between program completion and Anatomy and Physiology II

course grades, $r = .152$ and Microbiology course grades, $r = .191$, and between

NCLEX-RN® success and Anatomy and Physiology I course grades, $r = .171$.

Regarding preadmission test areas, Higgins (2005) found statistically significant

correlation coefficients between program completion and reading, $r = .124$, and between

NCLEX-RN® success and science, $r = .413$. No student demographic variables tested

were found to be statistically significant. The correlation coefficient was statistically

significant between NCLEX-RN® success and HESI Exit Examination scores, $r = .518$,

and the Nursing Skills Course grade, $r = .281$.

Phase 2 of the Higgins (2005) study involved interviewing other associate degree

nursing (ADN) program directors in Texas via telephone to determine strategies in use to

address attrition rates and NCLEX-RN® pass rates. Convenience sampling from a list of

Texas ADN programs resulted in 45 program directors providing verbal agreement for

participation. Four major themes of used strategies were identified through organization

of interview notes: preadmission requirements, campus counselors, remediation, and

faculty. Eleven programs reported using preadmission testing, some with specific scores

for admission eligibility. Other strategies included increased admission grade point

averages (GPA), increased prerequisite credit hours, limited admission numbers, and

restrictions on program reentry. Twelve of the directors interviewed mentioned using

campus counselors to address the attrition rate. Most of the directors voiced a form of

remediation; some hiring extra faculty for this task. Common criteria named for

remediation included examination scores below 70-75%, course grades of C or below,

one or early unit examination failure. Numerous forms of remediation were verbalized.

Information/data regarding raising the NCLEX-RN® pass rates were organized into four major themes: exit examinations, achievement testing throughout the curriculum, remediation, and adapting test items to reflect NCLEX-RN® style of testing. Thirty-three programs reported using exit examinations; HESI was the most commonly identified. Many of the program directors reported that achievement testing was in the beginning stages of implementation and use.

Phase 3 of the Higgins (2005) study consisted of telephone interviews with ten faculty at the research setting and 30 new graduates who had taken the NCLEX-RN® to elicit suggestions for lowering attrition rates and raising NCLEX-RN® pass rates. Randomized sampling was used to select interview candidates and verbal agreement was elicited to participate.

Faculty responses regarding attrition rates reflected three themes: prerequisites for admission, various forms of mentoring, and faculty needs. Faculty verbalized considerations of establishing cut-off scores for preadmission testing in reading, math, and verbal areas, and limiting prerequisite course repeats. Faculty needs included feeling overworked, which limited development of teaching strategies to effectively meet needs of students. Faculty stated that recognition of work achieved may motivate faculty. Student responses regarding attrition rates were uncertain; themes developed included the need for a mandatory class for test-taking skills, and encouraging use of ancillary study methods such as review books, test reviews, and study groups.

Higgins (2005) found that faculty responses regarding NCLEX-RN® pass rates fell into three themes: teaching, test-item writing, and curriculum changes. Faulty

reported the desire to include variation of teaching styles such as case studies, scenarios, content maps, and critical thinking exercises. Five faculty members (50%) suggested the need to revise the test bank and write more NCLEX-RN®-type questions. Faculty suggested changes to curriculum that focused on increasing content in pharmacology and medical-surgical areas, and aligning the two medical-surgical courses in consecutive order for students. Student responses for improving NCLEX-RN® pass rates were developed into four themes: the use of NCLEX-RN®-style testing, expanding class content, use of review books, and review of course tests. Student stated that prioritization and delegation were areas that needed increased curricula time; other suggested areas for concentration included pharmacology, communication skills, and critical thinking skills.

Strengths of Higgins' (2005) study include inclusion of methodology details to address trustworthiness of the qualitative data. Participant language, verbatim accounts, mechanically recorded data, member checking, and participant review were reported as used to strengthen credibility. Details were provided for sampling selection for each phase. For the quantitative data, the data collection sheet and coding process was described to address confidentiality. Weaknesses include that transferability will be limited, based on selecting student sample from three semesters, and the effect of passage of time. Dependability and confirmability were threatened since no audit trail was reported as used for this study. Instrument validity and reliability was not specifically addressed. A major weakness noted is the statistical analysis used. Results are reported as r values, which would indicate use of the Pearson product-moment correlation. This would be inappropriate to use, because both program completion and NCLEX-RN results

were stated as nominal data. Testing for the Phi coefficient would have been an appropriate statistic to use for this study.

Results from this Higgins (2005) study provide guidance for curriculum changes based on research findings. For this study sample, significant findings suggest the usefulness of the HESI Exit Examination for benchmarking.

Spurlock and Hunt (2008) conducted a retrospective descriptive, correlational study at a nursing program to investigate the disparity between the Health Education Systems, Incorporated (HESI) Exit Examination predictor scores and NCLEX-RN outcomes. Data collected were gathered from student records after both standardized exam instruments were completed. Convenience sampling was used based on complete and available student records, for a total sample size of 179 students. Students at this Midwestern nursing program were required to take the HESI Exit Exam until a score of 850 was achieved. Logistic regression analysis was used to predict NCLEX-RN outcomes from the Exit Exam scores. For testing the predictive nature of the HESI Exit Exam, the NCLEX-RN outcome was used as the independent variable, and one-way ANOVA was the descriptive statistic used to assess for differences in HESI Exit Exam first and final scores. Crosstabulation was used to examine the HESI Exit Exam scores of those students who passed versus those who failed the NCLEX-RN.

For Spurlock and Hunt's (2008) research question of: "What is the relationship between students' first Exit Exam scores and NCLEX-RN outcomes", findings revealed that multiple attempts to reach the benchmark score of 850 caused a decrease in relationship strength between the Exit Exam and the NCLEX-RN. There was a

statistically significant relationship found between the NCLEX-RN outcomes and first HESI Exit Exam score only, $r_{pb} = -0.275$, $p < 0.005$.

For Spurlock and Hunt's (2008) research question of: "Do HESI Exit Exam scores statistically significantly predict NCLEX-RN outcomes", first and final Exit Exam scores were run separately using binary logistic regression analysis. Findings revealed a "better-than-chance" distinguish between pass and fail for the first Exit Exam scores, and a poor prediction of NCLEX-RN failure. Final Exit Exam scores were found to be insignificant and poor predictors of NCLEX-RN outcomes.

For determining HESI Exit Exam cut-off scores for most accurate classification of NCLEX-RN pass versus fail, evaluation of data by logistic regression model revealed that the best HESI Exit Exam cutoff score for this sample was 650. This was based on first Exit Exam scores only, as the final scores were found to be insignificant predictively.

Spurlock and Hunt's (2008) final research question was "Do descriptors of HESI Exit Exam scores actually reflect the real probability of a student failing the NCLEX-RN?" Binary logistic regression model for the first Exit Exam scores was used. Findings for this sample revealed that the high-scoring categories were correct; students scoring in this category had little chance of NCLEX-RN failure, $p = 0.03$ to 0.05. The category of average probability of passing scores was incorrect; the average failure rate was actually higher than described by HESI. In the lowest scoring categories, considered at "grave risk" of fail and "poor performance expected", found actual NCLEX-RN testing risk to be at increased risk, but not as poor as predicted.

Findings of this program of research were used to evaluate and modify their progression policy and use of the HESI Exit Exam. While first exit Exam scores were found to be significant predictors of NCLEX-RN outcomes, it was noted that progression should not be determined on a single predictive test. Since only the first Exit Exam scores were significant for this sample, findings suggest that testing to achieve a specific score was not effective for NCLEX-RN success.

Strengths of Spurlock and Hunt's (2008) study include the inclusion of details of methodology and data collection. The human protection approval was reported and data collection strategies for confidentiality were adequately described. Weaknesses of this study include that instrument validity and reliability was not specifically addressed. This sample was homogenous in nature from a single program; therefore, the generalizability of the findings is limited. The type of nursing program, ADN, versus BSN, or diploma, used in this study was not specifically addressed.

Findings of Spurlock and Hunt's (2008) study emphasize the importance of using many variables to determine progression and benchmarking goals in nursing programs. While the HESI Exit Exam appears frequently in nursing literature as commonly used for nursing programs, and for good predictive nature based on numerous research findings, this study reveals different findings for its predictive nature. This demonstrates the importance of repeated research studies for each standardized exit examination in various program sizes, types, and locations for valuable results. Each nursing program needs to be able to use the most accurate, recent information from students in that specific

curriculum. Only then can these student data be used in the benchmarking process for that particular curriculum.

<div align="center">Other Variables of Interest</div>

Arathuzik and Aber (1998) conducted a descriptive, correlational study to identify factors associated with NCLEX-RN success at a public university nursing program with a diverse population. Convenience sampling was drawn from the senior students who were completing their final semester of curricula at an urban university ($N = 79$). Four instruments were used for this study: a demographic data sheet, the Internal Block Scale, the External Block Scale, and the Study Skills Self-Efficacy Instrument (SSSE). The dependent variable for this study was NCLEX-RN outcome.

Based on the Internal Block Scale, Arathuzik and Aber's (1998) findings revealed a variety of internal blocks to success in the nursing program, reported on a scale of 1 to 10; "very little" to "quite a lot", respectively. Means and significant deviations for these difficulties experienced during the nursing program included, respectively: emotional distress (5.18, 3.04), fatigue (6.53, 2.52), lack of confidence (5.19, 2.98), and multiple role strain (6.01, 3.05). The External Block Scale with the same Likert-type responses for effects experienced during the nursing program was used. Means and standard deviations for various external blocks identified included, respectively: finances (7.43, 3.63), family demands (5.78, 3.18), work demands (5.61, 3.11), and relationship strain (5.98, 3.03).

Arathuzik and Aber (1998) calculated point biserial correlation coefficients to test for relationships between the student data variables and NCLEX-RN outcome. At .05 level of significance, significant relationships were revealed between NCLEX-RN

outcome and cumulative undergraduate grade point average (GPA), $r_{pb} = .275$, English

as the primary language, $r_{pb} = .253$, and a sense of competency in taking tests that

necessitate critical thinking, $r_{pb} = .245$. At the same level of significance (.05), inverse

relationships were revealed between NCLEX-RN outcome and family demands or

responsibilities, $r_{pb} = -.293$, and emotional distress experienced, $r_{pb} = -.240$.

Strengths for Arathuzik and Aber's (1998) study include that instrument

reliability and validity measures for the Internal and External Block Scores and for the

SSSE were reported as adequate. Data collection measures for confidentiality were also

reported. Weaknesses include that generalization is limited due to small sample size and

single site research.

Findings from Arathuzik and Aber's (1998) study reveal the impact of internal

and external blocks on the learning process of this student sample. It is an important

reminder to assist students in identifying challenges or barriers to successful NCLEX-RN

testing, so that students can find strategies for dealing with these. As previous research

indicates, findings from this study revealed that higher GPA was significantly correlated

with NCLEX-RN results. There was a lack of evidence examining the relationship of

critical thinking abilities with NCLEX-RN outcomes.

Beeman and Waterhouse (2001) conducted an ex-post facto, quantitative study

using a discriminate analysis to identify predictors of NCLEX-RN success at a

baccalaureate nursing program in Delaware. Twenty-one independent variables of

interest were examined, including year of graduation, sex, age at graduation, traditional

versus accelerated nursing program, Scholastic Aptitude Test (SAT) verbal and math

scores, grades for two biology courses, grades for physiology and pathophysiology, grades for eight didactic nursing courses, number of C+ or lower grades received in nursing didactic courses, number of B or lower grades received in clinical courses, grade point average (GPA) for the end of the sophomore year, and GPA after the first senior semester. Instruments for this study included student records and the SAT.

Convenience sampling was used in the Beeman and Waterhouse (2001) study to collect computerized data from student records for the academic years 1995 through 1998; both traditional program and accelerated nursing students were included. The 538 participants were randomly divided into two equal groups. One group was used for calibration of the discriminant function. The other group was used to cross-validate the discriminant function later.

Findings by Beeman and Waterhouse (2001) show that NCLEX-RN success are significantly correlated with the number of C+ or lower grades in nursing didactic courses, $r = .384$, $p = .0001$, grades in two restorative nursing courses, $r = .379$, $- .381$, $p = .0001$, grades in the introductory nursing course, $r = .367$, and pathophysiology II course grades, $r = .349$. These variables were found to be significantly intercorrelated.

Beeman and Waterhouse (2001) reported that overall for this sample, 93.3 percent of students were correctly predicted for NCLEX-RN outcomes by the discriminant analysis: 92 percent accuracy for failing and 94 percent for successful testing. A Wilk's lambda for the NCLEX-RN pass and fail groups means was significant at the $p < .0001$ level for the discriminant analysis. There were significant differences between those passing and those failing the NCLEX-RN: those who passed held statistically significant

higher grades in all didactic nursing courses, had significantly higher GPA, and had a significantly lower number of low didactic (C+ or lower) and clinical (B or lower) grades. Significant differences were found between the program types, with fewer failures in the accelerated program graduates, $X^2 = 9.53$, $p = .013$. Also, younger nursing graduates were found more likely to fail the NCLEX-RN, $t = - 2.52$, $p = .002$.

Strengths of Beeman and Waterhouse's (2001) study include that it examined many readily available variables for predictive nature for NCLEX-RN results. The study provided adequate description of statistical analysis used, and adequate information regarding confidentiality of student records. Weaknesses include that generalizability was limited because of single site participation.

Beeman and Waterhouse's (2001) study was conducted shortly after adoption of the computerized adaptive testing format of the NCLEX-RN. Although the variables examined in this study do not include standardized, computerized assessment exams, findings demonstrate that discriminant function is a useful tool for timely identification of students at-risk for NCLEX-RN failure.

Haas et al. (2004) conducted a retrospective, correlative study at an upper-division nursing program in the southeastern United States to predict student success on the NCLEX-RN. Convenience sampling of existing student records was used for $N = 351$. Of those records, 34 indicated a failure on the first testing attempt on NCLEX-RN; the overall sample NCLEX-RN pass rate was 90.3% ($n = 317$). The independent variables studied for prediction in NCLEX-RN success were gender, race, age, nursing cumulative GPA, transfer undergraduate GPA, cumulative undergraduate GPA, verbal

and quantitative SAT scores, and campus location (main campus versus outlying campus). Merit scores were another variable used for comparison to other variables. These scores were assigned by the admissions office, based on a formula comprised of verbal SAT scores and GPA (Haas et al., 2004). Instruments used included student records, the SAT, and the NCLEX-RN.

Haas et al. (2004) analyzed data using SPSS software, setting the Type I alpha at .10, rather than the usual .05; this was done to potentially take a 10% chance of including the falsely identified failure students. Thus, values of $< .10$ were considered statistically significant for this sample. Discriminant analysis was used to predict groups based on variables, and to discover differences between the groups. When correlated to overall function, these identified significant predictors determined the hit rate, or the percentage of cases correctly predicted for classification to NCLEX-RN pass or fail.

Haas et al. (2004) used chi-square analysis to determine the relationship between NCLEX-RN success and gender or race. Men failed at a significantly higher rate than women, $p = .064$. Racial diversity also was found to be significant, $p = .011$. Pairwise comparisons of all possible pairs of racial groups were calculated in attempt to minimize effects of small sample size for Asian and Hispanic students. The fail rate for the African American group was significantly higher than that of the Caucasian group, $p = .056$, Fisher's exact test, two tailed. The Caucasian group passed NCLEX-RN at a significantly higher rate than the Asian group, $p = .026$, Fisher's exact test, two tailed. To determine effects of campus location on NCLEX-RN outcome, chi-square analysis was run with no significant differences revealed. The pass rate of students of the

outlying campus was 88.6% (n = 117); the pass rate of students of the main campus was 91.3% (n = 200).

Interval data comparisons were performed on the study variables (Haas, Nugent & Rule, 2004) to determine if significant differences existed between passers and failers. Variables of undergraduate GPA, nursing cumulative GPA, and merit scores were compared for significant differences. The only statistical difference found in these analyses was in nursing cumulative GPA, p = .000. This finding is in agreement with previous research findings. Those who passed NCLEX-RN were found to hold GPAs that were approximately .3 point higher than those who failed. Interval data of standardized exams and age differences comparing passers and failers were examined by using t-tests. Significant differences were revealed between passers and failers in both standardized measures, passers exceeding in verbal SAT, p = .001, and quantitative SAT, p = .082 scores. Passers were found to be significantly younger in age, p = .097.

Haas et al. (2004) used discriminant function analysis in a stepwise fashion to correctly identify two true groups of pass and fail for a hit rate of 100%. This created two other potential categories: "false passers", who were predicted to pass but failed, and "false failers", who were predicted to fail, but passed. Using the stepwise approach, predictors were added and removed until the strongest possible predictor was achieved. Additionally, the variables of race, campus location and gender were converted from nominal to ordinal data to account for small groups. Predictor variables used for this analysis were verbal and quantitative SAT scores, nursing cumulative GPA, age, race, campus location, and gender. The function created through use of these predictors was

found to be significant, $X^2 = 28.946$, $p = .000$. The overall hit rate was 70.1%, with

61.2% of true failers correctly identified. True passers were identified correctly 71.0% of

the time; false passers represented 29.0% of the hits, and false failers represented 38.7%.

These findings indicate that over 60% of the NCLEX-RN fail group were correctly

identified as failers using this discriminant function.

Strengths for the Haas et al. (2004) include appropriate statistical analysis

methodology for the variables used considering levels of measurement. The study design

is appropriate for data collected. Weaknesses include that generalizability is limited due

to single site research and small group sizes of Asian and Hispanic participants. The

report did not indicate what type of nursing program was used as the research setting,

so that impairs generalizability as well. Procedures for maintaining confidentiality in

data collection were not addressed. Instrument reliability and validity measures were

not provided.

Findings from this study (Haas et al., 2004) support the previous research

(Beeman & Waterhouse, 2001) revealing the useful nature of discriminant analysis in

identifying predictors of NCLEX-RN outcomes. In agreement with literature review

of research which explores predictors for NCLEX-RN, this study emphasizes the

importance of determining variables which can accurately identify student at-risk

of failure.

Yin and Burger (2003) utilized an ex-post facto study to correlate several

variables to success on the NCLEX-RN. Independent variables selected for correlation

included student age at admission; gender; race; type of student; high school cumulative

grade point average (GPA); high school class rank; ACT composite score; college GPA and number of college credit hours before entering nursing program; GPA at graduation; and performance on required pre-requisite courses. The data for this study were obtained retrospectively from one Midwestern associate nursing program at a state university from 1991 to 2001; the sample included 325 graduates. Instruments used were student records, the ACT, and NCLEX-RN. Logistic regression, t-test, and chi-square analyses were used to determine predictors.

Results of the Yin and Burger (2003) analyses found significant relationships with NCLEX-RN success: college GPA prior to admission to nursing program (M = 3.20, $SD = 0.38$), $t = 2.30$, $p = 0.024$; natural science GPAs ($M = 3.11$, $SD = 0.69$), $t = 2.45$, $p = 0.018$; course grade on introductory psychology ($M = 3.11$, $SD = 0.69$), $t = 3.09$, $p = 0.002$; and high school rank (M percentage = 13%, $SD = 7.34$); no significant t or chi-square). Logistic regression models were run to identify the best set of predictors for success on the NCLEX-RN. The overall likelihood ratio statistic for a two-variable model, college GPA prior to admission and high school rank, was significant: for each 0.1 increase of GPA, the odds of passing NCLEX-RN increased three-fold.

A strength of this study was that most of the variables of student data were available before or at admission selection to the nursing program (Yin & Burger, 2003). These identified predictive variables can guide selection criteria for admission and provide earlier identification of at-risk students for intervention. A weakness of this study was that findings of variables that were not statistically significant are inconsistent with previous study findings. A possible explanation for the inconsistent finding

regarding the variable of race is that the sample consisted of a limited number of non-White students; resulting in a small power for detecting significance.

Findings from Yin and Burger's (2003) study were helpful for consideration of variables of concern related to NCLEX-RN success. The findings demonstrated the need for selecting those factors that affect NLCLEX-RN success that programs can use in guiding curriculum. Although the Yin and Burger (2003) variables were not identical to the current investigation variables, it has become evident through research that early identification of at-risk students is beneficial to successful remediation plans. This is helpful for planning program benchmarks.

Giddens and Gloeckner (2005) conducted a nonexperimental, ex-post facto study to investigate the relationship of nursing student critical thinking skills performance on the NCLEX-RN at a university-based nursing program in the southwestern United States. Convenience sampling was used to collect variable data from existing student records from 1998 through 2001 for $N = 218$. Independent variables for this study included critical thinking skills and disposition, nursing GPA, gender and age. The California Critical Thinking Skills Test (CCTST) and the California Critical Thinking Disposition Inventory (CCTDI) were instruments used to measure the critical thinking variables. Other instruments for this study included student records and the NCLEX-RN. The CCTST and CCTDI were administered to the students at program start for entry data, and during the last semester of the nursing program as exit data.

Giddens and Gloeckner (2005) defined the CCTST as a measurement tool for critical thinking skills in college-age individuals, considered discipline-neutral. Six

subscale scores were provided from the CCTST: analysis, evaluation, inference, inductive reasoning, and deductive reasoning. The CCTDI was reported as an assessment tool for the extent which a participant possesses the disposition of the ideal critical thinker, measuring affective attitudinal dimensions of critical thinking (Giddens & Gloeckner, 2005). The CCTDI elicits Likert-type responses from testers; 75 statements on a 6-point scale ranging from "strongly agree" to "strongly disagree." The seven subscale scores provided for the CCTDI include truth-seeking, inquisitiveness, open-mindedness, confidence, analyticity, systematicity, and maturity.

Giddens and Gloeckner (2005) used discriminant analysis to determine the ability to correctly classify students for passing or failing NCLEX-RN based on CCTST and CCTDI scores. For this analysis, independent variables used were nursing GPA, CCTST, and CCTDI. Overall, findings revealed 92% of the students were correctly categorized for the NCLEX-RN outcome: 98% accurate for passing but 79% incorrectly classified for those who failed.

Giddens and Gloeckner (2005) evaluated age at graduation, gender, and nursing GPA to find differences between the pass and fail groups for comparison to past research findings. A t-test found nursing GPA significant to NCLEX-RN outcome for this sample ($n = 209$), $t = 5.3$, $p = .001$, $d = 1.38$.

An independent t-test was computed by Giddens and Gloeckner (2005) to compare the mean scores on the CCTST and CCTDI between the NCLEX-RN outcomes groups. On the entry CCTST, the pass group scores exceeded the fail group. The difference in means for three of the subscales ($n = 101$) was statistically significant,

with large effect sizes: total score, $t = 2.5$, $p = .015$, $d = 1.0$; analysis, $t = 2.4$, $p = .017$,

$d = 1.1$; and deductive reasoning, $t = 3.0$, $p = .003$, $d = 1.2$. On the exit CCTST, the pass

group had statistically significant higher mean scores on all six subscales ($n = 191$), with

a medium to large effect size. These results were: total score, $t = 3.0$, $p = .003$, $d = .81$;

analysis, $t = 2.2$, $p = .026$, $d = .60$; evaluation, $t = 2.2$, $p = .030$, $d = .59$; inference,

$t = 2.6$, $p = .011$, $d = .69$; inductive reasoning, $t = 2.7$, $p = .008$, $d = .72$; and deductive

reasoning, $t = 2.1$, $p = .035$, $d = .69$.

Giddens and Gloeckner (2005) found that independent t-tests of mean scores on

the entry CCTDI showed no significant differences between the pass and fail groups. On

the exit mean scores for the CCTDI ($n = 183$), findings revealed statistically significant

differences between the groups on five of the subscales: total score, $t = 2.6$, $p = .010$,

$d = .72$; truth-seeking, $t = 2.7$, $p = .007$, $d = .75$; open-mindedness, $t = 2.4$, $p = .015$,

$d = .64$; synthesis, $t = 2.2$, $p = .030$, $d = .60$; and maturity, $t = 3.6$, $p = .001$, $d = .96$.

Using a one-sample t-test, however, revealed no significant difference between the pass

and fail groups.

Giddens and Gloeckner (2005) used the entry and exit critical thinking scores

for pre- and post-paired comparison to evaluate for changes in critical thinking and

disposition. Findings revealed a significant changes on the CCTST deductive

reasoning subscale ($n = 84$): $t = 2.4$, $p = .02$, $d = .26$, and on the CCTDI confidence

subscale ($n = 84$): $t = 3.5$, $p = .001$, $d = .3$. The small effect size for this data may

decrease significance.

Strengths of this (Giddens & Gloeckner, 2005) study include identification of a gap in research examining critical thinking in correlation with NCLEX-RN outcomes. Data collection procedures to maintain confidentiality was provided. Findings of pronounced significant differences between pass and fail groups on exit scores support previous research. Weaknesses include that generalizability is limited due to single site research of BSN students. Though descriptions of the critical thinking instruments were provided, adequate information was not reported for CCTST and CCTDI instrument reliability and validity.

Findings of this study (Giddens & Gloeckner, 2005) indicate some growth in critical thinking skills through nursing curriculum for this sample. Nursing GPA was identified as statistically significant in relation to the NCLEX-RN outcome, which supports previous findings. Although these critical thinking instruments are not proposed variables for the currect investigation, this study emphasized the importance of identifying variables associated with NCLEX-RN performance.

Examining Student Satisfaction

Richards and Stone (2008) conducted a qualitative study at a large Midwestern baccalaureate nursing program to determine student response and satisfaction regarding use of standardized, comprehensive testing as an assessment tool. At this research setting, proctored testing occurred outside classroom time, beginning in the third semester with one test, and progressing to two or more standardized comprehensive exams in the last three semesters of the nursing program. Upon adoption of this testing tool for this program, use in the first semester was voluntary and free of charge for students;

participation was minimal. In the following semesters of the study, participation remained voluntary, but points toward classroom exams were available based on standardized testing scores. Remediation was provided and encouraged for those who did not achieve the benchmark score using a nonproctored test version.

Richards and Stone (2008) authored student surveys, which were completed during the spring 2006 semester, to collect feedback from students. Survey items included yes/no questions, Likert scale responses, and short answer responses. Survey questions areas related to testing barriers, type of assessment taken, benefits of testing, use of preparation materials, and recommendations for future use.

Surveyed data were entered into SPSS for analysis for the Richard and Stone (2008) study. Responses were reviewed and categorized into general summary areas. Responses were also summarized according to semester and analyzed for trends. Convenience sampling was used to conduct this study; 410 usable student surveys were returned for a response rate of 61.84%.

For this study (Richards & Stone, 2008) the most frequent test administered was the maternal/child test with $n = 125$; the least frequent test was pharmacology with $n = 69$. The highest number of returned survey was maternal/child ($n = 119$); the lowest number of returned surveys was in leadership ($n = 58$).

Six possible barrier areas were identified by Richards and Stone (2008): cost of program, time to test, unable to test on nonclass day, using the science lab for testing, schedule conflict, and problems with registration. The most frequent barrier reported by student was testing program cost ($n = 60$). Test time concerns were reported by 34

students and 10 students reported conflict with testing dates. Other barriers to testing reported by students included "too busy" and feelings that the testing format was "not useful". Forty-five percent of the students reported participation in the nonproctored assessments; 65% of those responding had scored at or above the benchmark score

Richards and Stone (2008) found that according to a Likert-type scale, 22 % of the students rated the testing and materials as very or somewhat useful. Twenty-seven percent rated the testing as neutral; 27 % rated the testing as somewhat useful or not useful. Twenty-four percent of the students did not respond to this question. Fifty-seven percent felt that the orientation provided for testing was not helpful.

Short answer responses were allowed for comments on further use of the testing program for Richards and Stone's (2008) study. Forty-seven students reported that the testing should remain optional. Twenty-five students felt that the testing should not affect their course grades, and 13 respondents felt the program should not be used for nursing students in the future.

Regarding cost and location of the testing program, findings by Richards and Stone (2008) revealed that eight students felt that the cost should be included in student tuition fees; 11 students stated that the testing cost was "too expensive." Five students recommended that the testing should occur only in the nursing building. Another issue raised by student responses was computer server problems while testing. Fifteen students reported on this issue and impacts of testing interruption.

Student feedback from the study (Richards & Stone, 2008) stimulated various changes to the testing program. It was anticipated that with continued use, students will

feel more comfortable and attitudes pertaining to testing will improve. Faculty documented plans to revise the testing orientation process by posting orientation materials within each class, adding a web page link, and communication via e-mail. Financial concerns were resolved by including testing fees into student fees and using financial aid to cover costs.

Strengths of Richards and Stone's (2008) study include an appropriate study design based on outcomes desired: student satisfaction regarding a new learning strategy. Weaknesses include that the actual test products used for this study are not identified. Also trustworthiness cannot be fully assessed because adequate information was not provided regarding member-checking, audit trails, number of researchers, size of research teams, or secondary review of information categories compiled. Transferability to other nursing program types and locations is limited because of single site research. The instruments used were not identified.

This study (Richards & Stone, 2008) was useful for the current investigation because it examined satisfaction with standardized computerized assessment testing in a nursing program. This study was unique because it focused on content mastery assessment scores only, and not end-of –program, or exit comprehensive testing as most literature reports. It is essential to remember that many factors surrounding standardized testing affect the student experience and outcome, including cost, schedule, location, and functional technology.

<center>Summary of Research</center>

Benchmarking as a means for quality improvement in nursing education has been documented in nursing literature, studied by identifying predictors and implemented in program progression policies (Spurlock, 2006). Much of the research conducted using standardized testing as a variable of interest has utilized the Health Education Systems, Incorporated (HESI) examinations. The HESI tests have been used at mid-curricular points for some programs, and/or at end of program points; both commonly as progression standards (Morrison, 2005). There remains a lack of documentation of research testing other standardized tests for benchmarks and their use in programs, and inadequate information regarding continued use and evaluation of tools. Continued exploration of useful tools with evaluation is essential to this ongoing process (Griffiths, Papastrat, Czekanski, & Hagan, 2004; Billings, 2007). What exactly is being benchmarked, how are benchmarks determined and evaluated, and how is that information used for the continuous improvement of nursing education?

An area of investigation regarding benchmarking is the use of standardized assessment tests. Utilizing standardized assessment testing of content areas in addition to the usual program curriculum and exams during nursing school can assure that students are achieving at levels which will lead to success on the NCLEX-RN® after graduation. Testing results can identify areas of student learning needs to individualize remediation plans, both during the program, and at program completion, before independent study toward NCLEX-RN® testing begins.

CHAPTER III-METHODOLOGY

Chapter I discussion suggested a need for the current investigation, while Chapter II reviewed the appropriate nursing literature. This chapter will discuss the methodology for the investigation.

Faculty at the research setting ADN program were interested in determining whether the student ATI (2005) content mastery assessment exam scores taken by third and fourth semester nursing students were statistically correlated with sub-scale scores on the comprehensive final standardized exam, the ATI RN-Predictor (2007c), taken in the final weeks of nursing school. Since each nursing curriculum is unique to that specific program, each school must investigate their own student variables.

Research Design

The research design selected for this investigation was a Level II (Brink & Wood, 1994) quantitative, non-experimental, correlational, ex-post facto design. There was no treatment or manipulation of variables. This level of research was appropriate since there was very little research at Level III and what has been conducted does not pertain to students at this particular school. The three variables of this investigation were nursing student archived ATI (2005) content mastery scores, ATI RN-Predictor scores and NCLEX-RN® outcome.

Selection of Sample

Convenience and retrospective sampling of nursing students of a Midwestern community college were used. Due to availability of student data and testing scores, convenience sampling was utilized for the purposes of this investigation, including two campuses using the same curriculum. Based on a power analysis using a formulation of

80% power, a medium critical effect size of 0.4 for each of the dependent variables, and a significant level of 0.5 for bivariate correlation, a sample of 50 was deemed sufficient (Polit & Hungler, 1999). Nursing student exam scores from academic years 2006-2007, 2007-2008, and 2008-2009 were used for a pool of 91 student scores. Students of all ages were included, as well as nursing students with other degrees, e.g., Associates of General Science. Students returning to the nursing program after failing a semester and students of the online-only format were eliminated from the sample. Only first-time ATI student scores were included in this investigation.

<div align="center">Protection of Human Participants</div>

Approval for this investigation was obtained from the Nursing Research Ethics Committee at Fort Hays State University (see Appendix A). Permission for this investigation was obtained from the community college Nursing Program Dean and the Vice President of Instruction at the community college (see Appendix B). There were no invasive or life-threatening procedures associated with this investigation.

The Office for Protection from Research Risks (OPRR), (1997, ¶ 14), under the auspice of the United State Department of Health and Human Services, exempts research activities that involve educational tests unless the participant(s) or their scores are identified. The OPRR also states that such exempt research investigations should have student participants that are neither "at risk of criminal or civil liability" nor have risks to "financial standing, employability, or reputation---45CFR46.101" (¶ 14). The findings of this research will be used by the community college research setting for curriculum

clarification, modification or alteration. Therefore, individual students/graduates were not required to sign for permission to be in the investigation.

To further protect the participants, all data were reported in aggregate form. There were no identified risks to student participants. Nursing student progression in the major was not hindered or changed; most of these students had already graduated from the nursing major. There were no known benefits for the student participants; however, the community college nursing curriculum can gain valuable information for curriculum adjustment and was therefore identified as a substantial benefit for the program of nursing. Archived data collected did not contain student identifiers.

Data Collection

Student academic data, including ATI (2005 & 2007c) testing scores, and NCLEX-RN® results were obtained anonymously from existing students records kept at the nursing program. A research assistant, approved by the community college Department of Nursing, gathered the data. Each set of student data was coded with an assigned number to provide anonymity of specific student information. Coded data were transcribed to a form for computation (see Appendix C). Codes were known only to the research setting and the research assistant, not the principal investigator. Data were entered into PASW®, version 18 (SPSS, 2010). Copies of all data files were provided to the research setting for their own tracking and trending purposes for program national accreditation and statistical approval processes.

Instruments

There were three variables in this investigation. Each variable will be described and reviewed.

ATI Content Mastery Series 2.1 Scores

The ATI (2007d) utilizes multiple sources of evidence to assure the highest degree of validity for testing. Content Mastery (CM) exam items were based on the most recent detailed NCLEX-RN® test plan available and were administered in proctored settings and online format. Initial ATI construct validity used teacher and student ratings of test purpose and utility in applied test settings. Analysis of data during the developmental phase resulted in a coefficient alpha average of .93, and student ratings averaged 0.72 (ATI, 2007d). Individual scores were provided in percentage correct for each test.

Content Mastery of RN Adult Medical-Surgical Nursing (CM-MS) was a 90-item test to assess student's basic comprehension and mastery of adult medical and surgical principles. The 90 items were asked in multiple-choice fashion with primarily four choices available. The possible range of scores is 0-100%. The national mean was currently 67.4. The CM-MS was focused on principles including: the nursing process, therapeutic communication client education, cultural competence, pharmacology, nutrition, and holistic care for health promotion and disease prevention (ATI, 2007a). The ATI CM-MS had current Cronbach alpha reliability coefficient of .695.

Content Mastery of RN Maternal-Newborn Nursing (CM-MN) was a 60-item exam, assessing basic comprehension and mastery of maternal-newborn nursing

principles. These exam items were also primarily multiple-choice with four possible answers. The range of possible scores was 0-100%. The national mean for the CM-MN was 65.6. The CM-MN assessed principles including appropriate pharmacologic and nutritional interventions and cultural considerations (ATI, 2007a). The current CM-MN Cronbach alpha reliability coefficient was .682.

Content Mastery of RN Nursing Care of Children (CM-NC) was a 60-item test assessing basic comprehension and mastery of principles essential to quality nursing care of children. The items on the test were primarily asked in multiple-choice question fashion with four possible selections. The range of possible scores was 0-100%. The national mean was currently 66.8. The concepts covered in the CM-NC include the following: basic care, system disorder, pediatric emergencies, and psychosocial development disorders (ATI, 2007a). The current CM-NC Cronbach alpha reliability coefficient was .646.

Content Mastery of RN Mental Health Nursing (CM-MHN) was a 60-item test assessing basic comprehension and mastery of mental health nursing principles. The test items were asked in multiple-choice format with primarily four possible answers. The possible range for the CM-MHN was 0-100%. The national mean for the CM-MHN was 65.2. Concepts included on the CM-MHN were basic care, non-pharmacologic therapy, pharmacologic therapy, and care of complex mental health disorders (ATI, 2007a). The current CM-MHN Cronbach alpha reliability was .657.

ATI RN-Predictor®3.0

The 3.0 RN-Predictor® comprehensive exam (ATI, 2007c) was a 180-item exam based on the 2007 NCLEX-RN® detailed test plan for the purpose of assessing a student's comprehensive nursing knowledge base prior to graduation. Item format included 155 traditional, four option, multiple-choice questions with a single correct response. Additionally, the 3.0 RN-Predictor® comprehensive exam contained 25 alternate-format items of three types: multiple-response, fill-in-the-blank, and hot spot. Sub-Scale area scores of interest for comparison in this investigation include four of the major content areas: Maternal-Newborn Nursing (P-MN), Adult Medical-Surgical Nursing (P-MS), Nursing Care of Children (P-NC), and Mental Health Nursing (P-MHN). The other major content areas assessed on the 3.0 RN-Predictor® comprehensive exam were Community Health Nursing and Leadership Nursing. Individualized student diagnostic results generated after testing includes: individual participant scores in percentage correct; predicted probability of passing NCLEX-RN®; individual scores in percentage correct within sub-scale topic categories, and within NCLEX-RN® outcome needs; directed study plan for individual review and remediation; and national and program means and percentile ranks (ATI, 2007c). The possible range of individual composite scores was 0 – 100%. The national mean was currently 64.7. Item analysis by ATI of the 3.0 RN-Predictor® resulted in a mean item difficulty of .65, with a mean point-biserial discrimination value of .22. Reliability by ATI using coefficient alpha resulted in an overall score of .75.

Plan for Statistical Analysis

The PASW®, version 18 (SPSS) was used to compute the statistics. Each research question was addressed separately.

Research question number one was, "Are there statistically significant relationships between students' ATI Content Mastery Series™ scores and ATI RN-Predictor sub-scale scores in the like content areas?" Student test scores were scale level of measurement. For research question number one, Pearson correlation coefficients were used to explore associations between each identified ATI content area score and the sub-scale score on the RN-Predictor exam.

Research question number two was, "Is there a statistically significant relationship between the student ATI RN-Predictor Comprehensive score and NCLEX-RN® outcome?" For research question number two, point bi-serial correlational coefficients would be appropriate to explore associations between the ATI RN-Predictor score and NCLEX-RN® outcome. However, SPSS does not offer this test so the Spearman rho was used. Comprehensive standardized exam scores were scale level of measurement. The dependent variable for this research question, NCLEX-RN® outcome, was nominal level of measurement, pass or fail.

Research question number three was, "Are there statistically significant relationships between students' ATI Content Mastery Series™ scores and NCLEX-RN® outcome?" For research question number three, Spearman rho correlation coefficient calculations were used to explore the associations between the Content Mastery scores of interest and the NCLEX-RN® outcome.

During data analysis, an ad hoc research question was formed for analysis: "Are there statistically significant relationships between the students' RN-Predictor sub-scales scores and NCLEX-RN® outcome?" For this research question, Spearman rho correlation coefficient calculations were used to explore the associations between these variables. Student assessment scores were scale level student scores while the NCLEX-RN pass or fail was nominal level of measurement.

<div align="center">Summary</div>

Standardized assessment testing can identify student content areas below proficiency level and provide remediation plans for individuals before program completion. Faculty at this nursing program can then set a benchmark testing score according to program and individual needs. Correlation of content mastery tests with sub-scale areas of Adult Medical-Surgical Nursing, Maternal-Newborn Nursing, Nursing Care of Children, and Mental Health Nursing on the comprehensive predictive exam can guide benchmark scores, as well as guide curriculum for the program. Further analysis of correlation between the comprehensive exam and the NCLEX-RN® outcome can determine usefulness for final student preparation and remediation. Collection and analysis of assessment testing data will be useful for continued tracking and trending of curriculum needs of this program.

CHAPTER IV – FINDINGS

This investigation examined the relationships between ATI assessment scores, comparing the Content Mastery™ Assessment scores to the ATI RN-Predictor scores in the like areas of Medical-Surgical, Maternal-Newborn, Nursing Care of Children, and Mental Health Nursing. The relationship between the ATI RN-Predictor comprehensive score and NCLEX-RN® outcome was also analyzed.

This chapter presents the findings of the data that were collected and analyzed from a Midwestern Associate Degree Nursing (ADN) program from student records from the academic years 2006-2007, 2007-2008, and 2008-2009. A research assistant collected the data from existing student records, student data were coded with anonymous identification numbers. Data were transferred from the original data form to the Predictive Analytics Software (PASW®) version 18 for analysis. The level of significance for this bivariate correlational investigation was set at 0.5.

Sample Characteristics

Student ATI scores, demographic data and first-attempt NCLEX-RN® outcome, pass or fail, data were gathered by a research assistant from three years of graduates from an ADN program at two campuses. The research assistant eliminated records which did not meet the delimitations developed for this investigation. Only those records which met all inclusive criteria were compiled in the data collection worksheet (see Appendix C). This resulted in a sample size of 91 participants with complete data for analysis. Participant demographic data gathered included age of student at graduation, gender, and

ethnicity. Of the investigation sample, 91.2% (*n* = 83) of the students were Caucasian,

6.6% (*n* = 6) Black, and 2.2% (*n* = 2) Hispanic. Graduate participant gender comprised

of 91.2% (*n* = 83) females and 8.8% (*n* = 8) males. This sample included students

from two campuses, 60.4% (*n* = 55) attended course on the main campus, and 39.6%

(*n* = 36) attended courses at the branch campus. The largest group (*n* = 41, 45%) of the

student sample was between 22 to 30 years of age at graduation; 22% (*n = 20)* of the

sample were 21 years or younger, and 33% (*n* = 30) were older than 30 years. The

NCLEX-RN® first-attempt student outcome for this sample, between years 2006-2009,

was 87.9% (*n* = 80) that passed on the first-attempt and 12.1% (*n* = 11) that failed.

Findings of Research Questions

Three research questions guided this investigation, and an additional research

question was explored during analysis. Each question will be individually addressed.

Research Question Number One

Research question number one was, "Are there statistically significant

relationships between student ATI™ Content Mastery Series scores and ATI™

RN-Predictor sub-scale scores in the like content areas?" Student records reporting

Content Mastery Maternal-Newborn (CM-MN) scores were low (*n* = 5); therefore, this

relationship was not examined. Thus, this research question included the Content

Mastery scores, taken after completion of didactic content during curriculum and

RN- Predictor scores in the following areas: medical-surgical, children, and mental

health. Medical-Surgical Content Mastery (CM-MS) (*M* = 65.36. *SD* = 8.77,

skewness = -.637, kurtosis = 1.309) was correlated with Medical-Surgical RN- Predictor

(P-MS) (*M* = 53.7, *SD* = 8.8, skewness = .488, kurtosis = -.249) and a weak, direct, statistically significant correlation was found, *r*(89) = .272, *p* = 0.009 (see Table 1).

Content Mastery student scores on Nursing Care of Children (CM-NC) (*M* = 63.3, *SD* = 9.24, skewness = -.244, kurtosis = -.152) was correlated with Nursing Care of Children RN Predictor scores (P-NC) (*M* = 70.33, *SD* = 11.16, skewness = .024, kurtosis = -.359). Using the Pearson *r* the results identified a moderate, direct, significant correlation, *r*(89) = .347, *p* = 0.001.

Content Mastery scores on Mental Health (CM-MHN) (*M* = 64.25, *SD* = 7.70, skewness = .177, kurtosis = -.098) was correlated with ATI RN-Predictor score on the RN Predictor Mental Health (P-MHN) (*M* = 70.31, *SD* = 13.92, skewness = -.205, kurtosis = -.728). A Pearson's correlation coefficients analysis was calculated which revealed the following direct, low, non-significant relationship, *r*(89) = .180, *p* = .08. See Table 1 for a summary of this data.

Research Question Number Two

Research question number two was, "Is there a statistically significant relationship between the student ATI RN-Predictor® Comprehensive score and NCLEX-RN® outcome?" The student ATI RN-Predictor (ATI-RN) total scores (*M* = 71.04, *SD* = 6.58, skewness = -.116, kurtosis = -.657) were correlated significantly with NCLEX-RN results, pass or fail, using Spearman's rho, *r*$_s$ (89) = .231, *p* = 0.028 (see Table 2).

Table 1

Pearson's Correlation Coefficients (two-tailed) of ATI Assessment Scores
(N = 91)

Assessment	*M*	*SD*	*r*	*p*
Medical-Surgical			.272*	.009
Content Mastery	65.36	8.77		
RN-Predictor	69.70	8.80		
Nursing Care of Children			.347**	.001
Content Mastery	63.31	9.24		
RN-Predictor	70.33	11.15		
Mental Health			.180	.088
Content Mastery	64.25	7.70		
RN-Predictor	70.31	13.92		

Note. $*p < .01.$ $**p < .001.$

Research Question Number Three

Research question number three was, "Are there statistically significant relationships between student ATI Content Mastery Series™ scores and NCLEX-RN® outcome?" Spearman's rho was used to calculate these correlations. Between the student CM-MS scores and the NCLEX-RN® outcome a direct, moderate, statistically significant relationship was identified, $r_s (89) = .348, p = 0.001$.

For the relationship between the student scores on the CM-NC and NCLEX-RN® outcome, a Spearman rho was calculated, $r_s (89) = .243, p = .021$. A weak, direct correlation that was statistically significant was identified.

Table 2

Spearman's Rho Correlation Coefficients (two-tailed) of ATI™ Scores and NCLEX-RN® Outcome (N = 91)

ATI Score	r	p
RN-Predictor Comprehensive score	.231*	.028
Content Mastery Medical-Surgical	.348**	.001
Content Mastery Nursing Care of Children	.243*	.021
Content Mastery Mental Health	.091	.393
RN-Predictor Medical-Surgical	.073	.494
RN-Predictor Maternal-Newborn	.184	.081
RN-Predictor Nursing Care of Children	.207*	.049
RN-Predictor Mental Health	-.005	.966

*p <.05, **p < .01*

Between the participant CM-MHN scores and NCLEX-RN® outcome the following Spearman's rho was calculated, $r_s(89) = .091$, $p = .393$. These findings show a relationship that was extremely weak, direct, and non-significant. See Table 2 for a concise report of the analyzed data results for research question number three.

Ad Hoc Research Question

An additional research query was identified as relevant for this investigation: "Is there a statistically significantly relationship between ATI RN-Predictor Comprehensive sub-scale scores and NCLEX-RN® outcome?" Spearman's rho correlation coefficients were used to analyze these relationships. The relationship between student P-MS scores

and the NCLEX-RN® were calculated with the following results, $r_s (89) = .073, p = .494$. This correlation was direct but extremely weak and non-significant.

The correlation between the students ATI RN-Predictor Maternal (P-MN) and NCLEX-RN® was calculated, $r_s (89) = .184, p = .081$. The results showed a weak, direct and non-significant relationship.

Between the P-MHN participants scores and the NCLEX-RN® results a Spearman rho was computed, $r_s (89) = -.005, p = .966$. The results showed an extremely weak, indirect, non-significant relationship.

A statistically significant, weak, and direct correlation was found between P-NC student scores and NCLEX-RN® outcome, $r_s (89) = .207, p = .049$.

Summary

This chapter has presented a description of the participant sample for the purposed investigation in CHAPTERs I and III. In addition, the findings for each research question for this investigation was reported. The interpretation of the investigation findings will be discussed in CHAPTER V.

CHAPTER V – SUMMARY AND CONCLUSIONS

This chapter provides a summary of the investigation interpretation of the findings and discussion of conclusions. Limitations of the investigation will be discussed along with recommendations for future research. Interpretations of the findings include the limited ability to generalize to all populations.

Summary of the Investigation

This investigation examined ATI™ Content Mastery assessment and RN-Predictor scores to determine if significant correlations existed between student scores on like content and between ATI student scores and NCLEX-RN® outcome. The investigation site was a rural Midwestern community college with an Associate Degree Nursing (ADN) program of two campuses. A convenience sample of 91 student records meeting inclusive criteria was used from the academic years of 2006-2009.

Interpretation of the Findings

Interpretation of data analysis findings will be included in this chapter. These findings are compared to studies in the nursing literature base.

Research Question Number One

Research question number one was, "Are there statistically significant relationships between student ATI™ Content Mastery Series scores and ATI™ RN-Predictor sub-scale scores in the like content areas?" Pearson correlation coefficients were calculated for these variables.

A weak, direct, significant correlation between the two examinations covering medical-surgical content, $r(89) = .272, p = .009$, was found; a significant linear

relationship between assessment testing of this content after didactic learning and at end-of-program curriculum exists. This finding indicates that participants that scored well in the Medical-Surgical content examination area during the nursing program were slightly more inclined to score well on the RN-Predictor examination in this content.

The mean score for the CM-MS for the investigation sample of students' scores in the year 2006-2007 ($n = 31$) was 64.08, while the national ADN program CM-MS mean score was 68.6. For 2007-2008, the investigation group ($n = 32$) score mean was 70.26, with the national standardized ADN program mean of 67.9. For the 2008-2009, an updated test version (CM 2007, Form B) was used. For this investigation group ($n = 28$), the mean score was ($M = 61.20$, $SD = 7.23$), with the national ADN program mean of 61.9. For the investigation sample ($N = 91$), score mean was ($M = 65.37$, $SD = 8.77$) over the three years (see Figure 2).

The mean score for the P-MS for the investigation sample for 2006-2007 ($n = 31$) was 63.71, while the national ADN program score mean (version 3.0) was 64.3. For the sample for 2007-2008 ($n = 32$), the P-MS score mean was 72.55. For the 2008-2009 sample ($n = 28$), the mean for the P-MS was 73.08. The reported ATI P-MS (version 2007B) was 69.8, while the computed score mean for the three-year sample was ($M = 69.70$, $SD = 8.80$). See Figure 2 for assessment score mean comparisons.

A moderate, direct and significant correlation was found between the CM-NC and the P-NC, $r(89) = .347$, $p = .001$. This indicates that participants that scored well in the CM-NC content during the nursing program were more likely to score well in the like content on the P-NC.
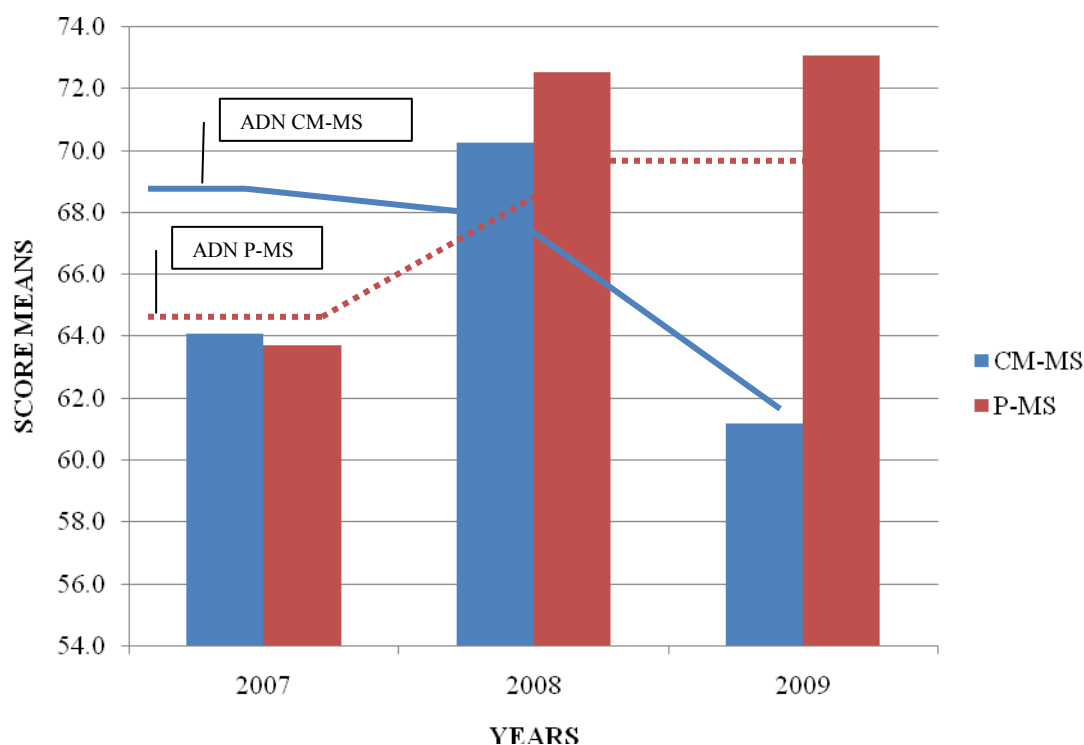
Figure 2. Medical-Surgical Score Means

The mean score for the CM-NC for the investigation group in 2006-2007 was 66.02, while the national CM-NC ADN program mean was 70.6. For 2007-2008, the investigation student group mean score was 66.09, with the national ADN CM-NC program mean score of 70.0. For the 2008-2009 group, an updated test version (CM, 2007, Form B) was used. For this group, the CM-NC investigation mean score was 57.15, with the national CM-NC ADN program mean score 62.5. Over the three years of research, the sample ($N = 91$), CM-NC score mean was ($M = 63.31$, $SD = 9.24$).

The mean score for the P-NC for the investigation sample for 2006-2007 ($n = 31$) was 71.29, while the ADN program score mean (version 3.0) was 72.4. For the sample

for 2007-2008 (*n* = 32), the national P-NC score mean was 70.31. For the 2008-2009 sample (*n* = 28), the mean for the P-NC was 69.28. The reported ATI P-NC (version 2007B) mean was 73.1, while the computed score mean for the three-year sample was (*M* = 70.33, *SD* = 11.16). See Figure 3 for assessment score mean comparisons.

A Pearson correlation coefficient calculated to examine the relationship between CM-MHN and like-content on the RN-Predictor found a weak, direct correlation that was not significant, *r*(89) = .180, *p* = .088. This finding indicates that the scores after didactic learning and at end-of-program for Mental Health have a non-significant relationship (see Figure 4).

It would be expected for students to achieve well in same content areas during nursing school and at the end before attempting NCLEX-RN® testing. With identified areas of improvement, it would also be expected for scores to improve from Content Mastery testing to RN-Predictor testing. Lack of significant correlation may suggest poor student remediation or preparation before RN-Predictor testing. Another possible influence is the time factor between didactic learning earlier in the program and knowledge loss. It was suggested by the faculty at the investigation site that students have verbalized lack of effort on the Content Mastery assessment testing because of time constraints. Scheduling issues and time to test were major barriers identified by a study (Richards & Stone, 2008) which explored student satisfaction of computerized standardized assessments. Findings from this research question provide guidance for future research and/or future use of this testing at this site.
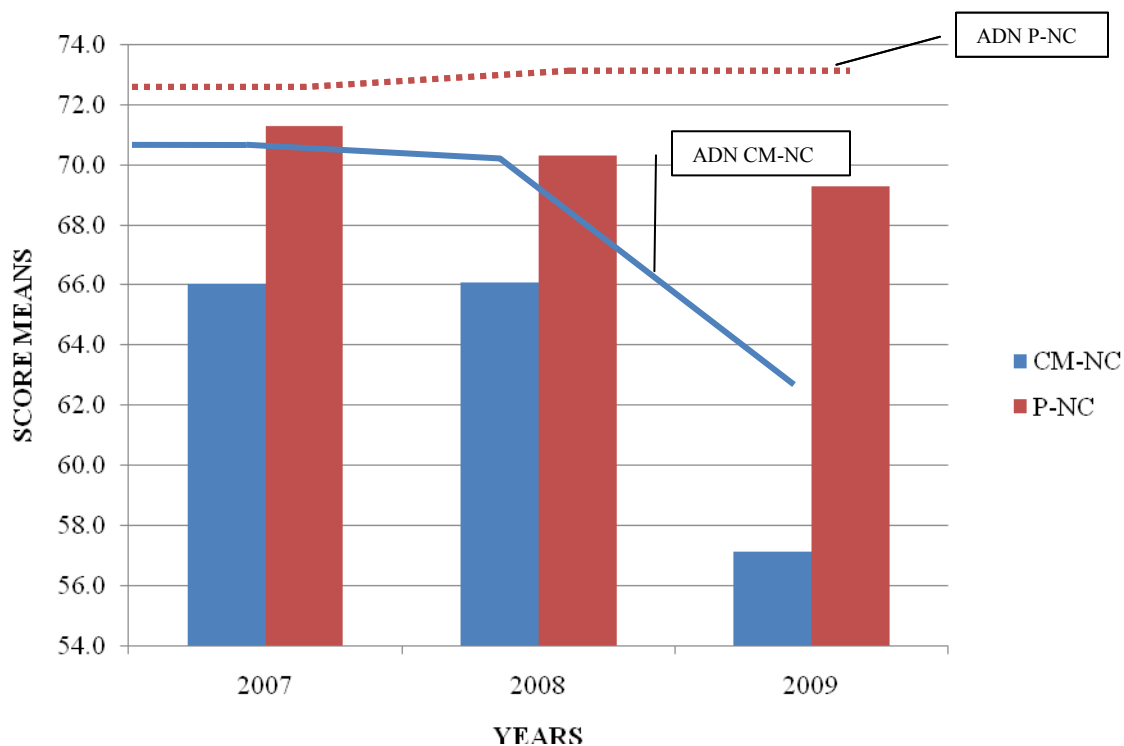
Figure 3. Nursing Care of Children Score Means

The mean score for the CM-MHN for the investigation group in 2006-2007 was

62.09, while the national CM-MHN ADN program mean score was 67.1. For 2007-2008,

the students' group mean score was 64.0, with the national CM-MHN ADN program

mean score of 67.1. For the 2008-2009, an updated test version (CM 2007, Form B) was

used. For this student group , the mean score was 66.93, with the national CM-MHN

ADN program mean score of 70.4. For the whole sample ($N = 91$), the score mean was

($M = 64.25$, $SD = 8.77$). See Figure 4 for assessment score means by academic years.
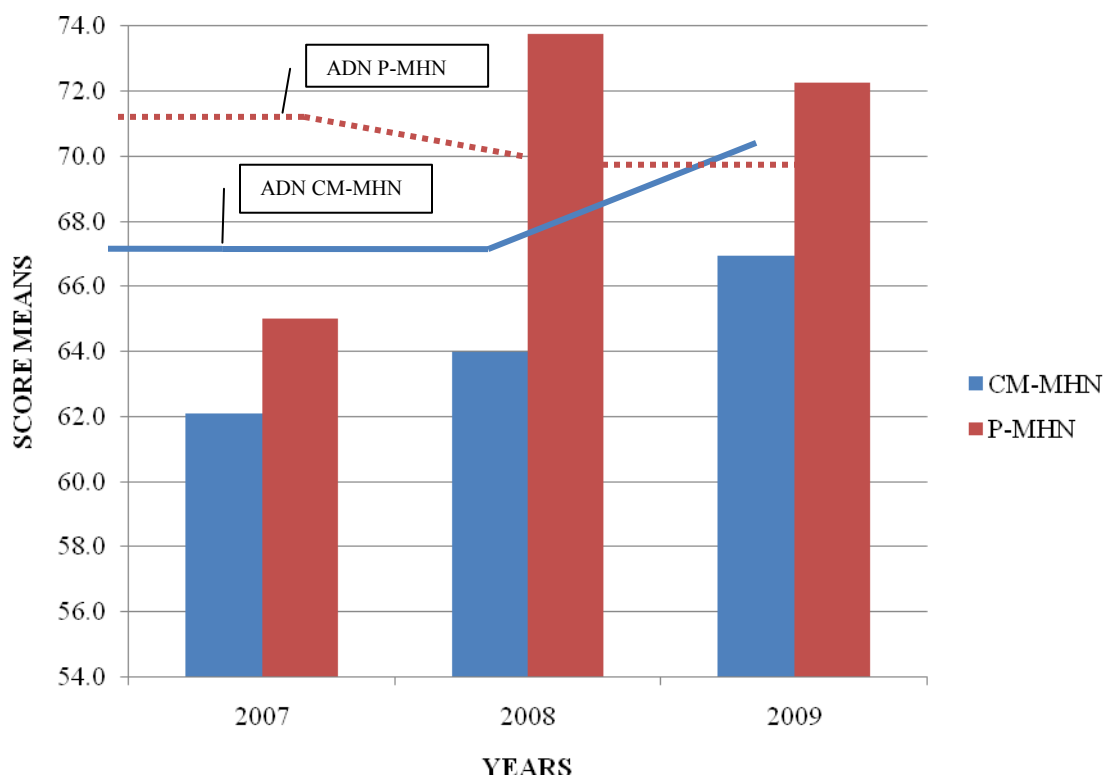
Figure 4. Mental Health Nursing Score Means

The mean score for the P-MHN for the investigation sample for 2006-2007 ($n = 31$) was 65.00, while the ADN program national score mean (version 3.0) was 71.3. For the sample for 2007-2008 ($n = 32$), the P-MHN score mean was 73.75. For the 2008-2009 sample ($n = 28$), the mean for the P-MHN was 72.26. The reported ATI P-MHN (version 2007B) mean was 69.6, while the computed score mean for the three-year sample was ($M = 70.31$, $SD = 13.92$). See Figure 4 for assessment score mean comparisons. Studies by Alexander and Brophy (1997) and Uyehara et al. (2007) concur with this finding of nonsignificance for Mental Health course grades. Further analysis by

Alexander and Brophy found an accurate predictive ability of 80.63% for NCLEX-RN®

when using logistic regression analysis of NLN comprehensive score in combination with

the Mental Health course grade.

Few nursing studies have examined correlations between standardized

mid-curriculum assessment testing scores and end-of-program scores in like-content

areas. A doctoral dissertation (Carl, 2007) explored ATI test scores as a predictor of

first time NCLEX-RN® outcome using Pearson product-moment correlation (2-tailed).

Findings indicated that six out of eight scores (RN-Predictor [2.0], TEAS entrance test,

Content Mastery series Leadership, Critical thinking, English, Reading and Science)

were statistically significant in correlation with NCLEX-RN® outcome. However,

only three means of test scores met the desirable .80 power analysis level: TEAS,

r = .411, p < 0.01, Critical Thinking, r = .427, p < 0.01, and English, r = .466, p < 0.01.

Actual *p*-values were not reported in the findings. The ATI scores used for the study

by Carl (2007) did not include the same Content Mastery score variables used for

this current investigation. Therefore, comparison of findings for this research question

was not possible.

This investigation was unable to analyze the correlation between the student

CM-MN mean scores and P-MN scores due to low number of student reported scores for

the CM-MN. The mean score for the P-MN for the investigation sample for 2006-2007

(*n* = 31) was 58.39, while the ADN program national score mean (version 3.0) was 55.0.

For the investigation sample for 2007-2008 (*n* = 32), the P-MN score mean was 68.48.

For the 2008-2009 sample (*n* = 28), the mean for the P-MN was 71.73. The reported ATI

P-MN (version 2007B) mean for ADN programs who took the examination in the nation

was 69.7, while the computed score mean for the three-year sample was ($M = 66.04$,

$SD = 12.34$). See Figure 5 for P-MN for this sample.

<div align="center">Research Question Number Two</div>

Research question number two was, "Is there a statistically significant

relationship between the student ATI RN-Predictor® Comprehensive score and

NCLEX-RN® outcome?" A Spearman rho correlation coefficient was calculated for the

relationship between these variables. A weak, positive, significant correlation was found,

$r_s(89) = .231$, $p = .028$. This significant correlation between the ATI RN-Predictor and

NCLEX-RN outcome suggests that students that achieved higher scores on the

RN-Predictor were more likely to succeed on first-attempt NCLEX-RN® testing.
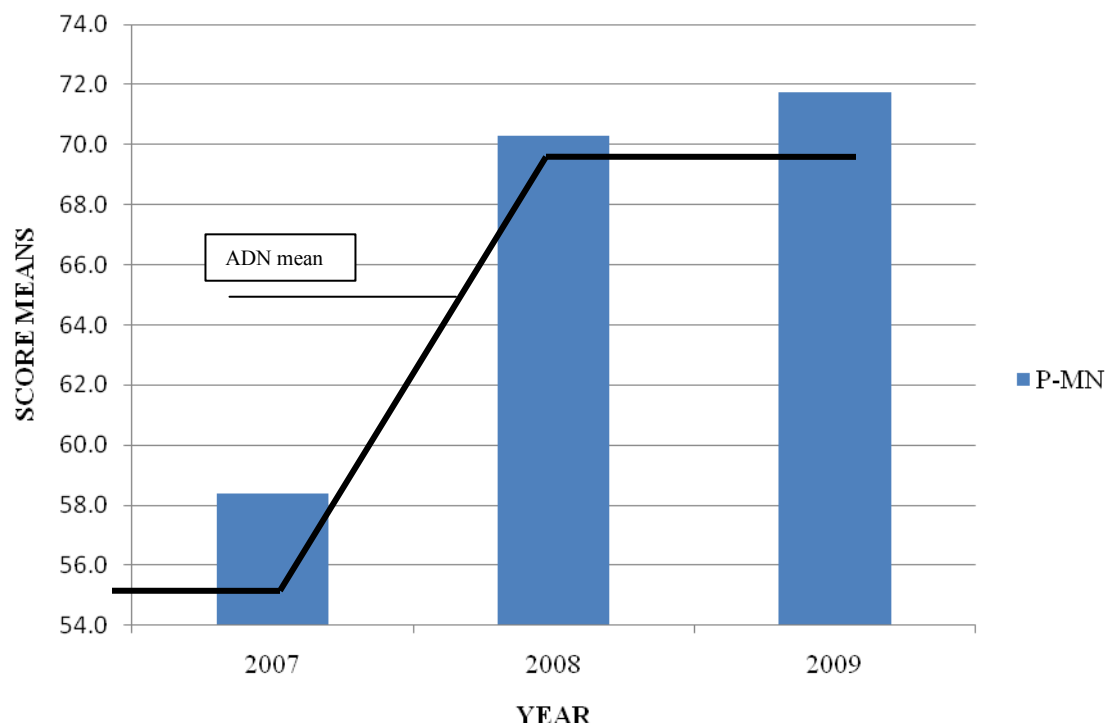
However, the correlation was weak.

Figure 5. RN-Predictor Maternal Newborn Care Score Means

For the sample in 2006-2007, the students' RN-Predictor mean score was 65.00; the national ADN program mean score was 64.7. An updated RN-Predictor was administered starting with the 2007-2008 sample. The mean score for the sample students on the RN-Predictor 2007 was 74.33, with a national ADN program score mean of 70.2. For the 2008-2009 student sample, the mean score was 73.96. The national ADN program mean score was 72.5. The students from the investigation program were above the national ADN score mean for the RN-Predictor for all of the studied years (see Figure 6).
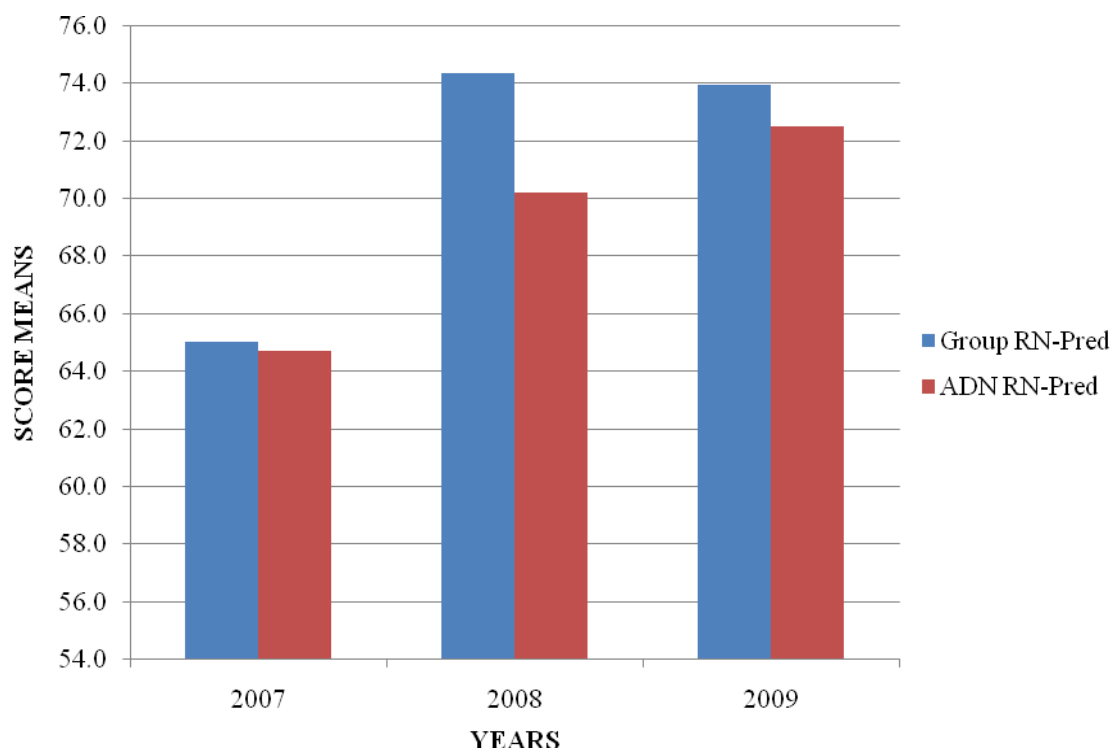
Figure 6. RN-Predictor Sample Score Means and National Means for ADN Programs

Nursing studies (Alexander & Brophy, 1997; Briscoe & Anema, 1999; Campbell & Dickson, 1996; Uyehara et al., 2007; Washington & Perkel, 2001) have reported similar significant correlations using various standardized student exit-assessment testing. Numerous studies have evaluated the HESI Exit assessment with significant predictive findings (Daley et al., 2003; Daigle, 2007; Higgins, 2005; Morrison et al., 2002; Nibert, et al., 2003; Rhodes, 2006; Spurlock & Hunt, 2008). The study by Rhodes (2006) used Spearman's rho to evaluate correlation of the HESI E^2, finding a moderate direct correlation, $r_s(75) = .48$, $p = .000$.

A correlational study (Daigle, 2007) conducted with an ADN sample at a similar setting as this current investigation used both the ATI RN-Predictor and the HESI E^2 exit exam to evaluate correlations with the NCLEX-RN outcome. Analysis using Spearman rho found significant correlations with both exams, the ATI, r_s (33) = .536, p = 000, and the HESI E^2, r_s (104) = .320, p = .000. This study reports moderate direct correlations with both exams, with the ATI of higher mean value than the HESI E^2. However, the sample number for the ATI was low, resulting in a skewness = -.391 and kurtosis = 19.72 for this finding.

Correlation studies specific to the ATI RN-Predictor with ADN student participants included studies by Carl (2007) and Kline (2010). The study by Carl evaluated the ATI RN-Predictor version 2.0. Though found statistically significant using Pearson product-moment correlation (2-tailed), the RN-Predictor did not meet the desirable .80 power analysis. Kline's (2010) study used logistic regression analysis and found strong predictive value with the RN-Predictor. However, the ATI test version used was not clearly identified in the report.

Findings from this research question indicate that the ATI RN-Predictor student scores are useful for determining students' outcomes on first-attempt NCLEX-RN® testing at this investigation school of nursing with this curriculum and is a beneficial tool for the benchmarking process. Generalizability of this finding to other programs is limited due to geographical location and is based on this site's curriculum.

Research Question Number Three

Research question number three was, "Are there statistically significant relationships between student ATI Content Mastery Series™ scores and NCLEX-RN® outcome?" Spearman rho correlation coefficient was calculated for relationships between these variables.

Between Content Mastery Medical-Surgical (CM-MS) scores and NCLEX-RN® outcome, a moderate, direct significant correlation was found, $r_s(89) = .348$, $p = .001$. This finding indicates that students that scored well on the CM-MS, taken at completion of didactic learning at this program, were more inclined to be successful on first-attempt NCLEX-RN® testing. For benchmarking purposes, this score may be useful for evaluating mastery of this content for this program, as well as, identifying at-risk students.

Similar results were found in the study by Uyehara et al. (2007) that reported significant correlations when comparing National League for Nursing (NLN) (year/version of the examination were not reported) standardized comprehensive testing content areas to NCLEX-RN® outcome. Uyehara et al. (2007) reported that content areas with significant correlations included NLN Adult Health Comprehensive, NLN Pediatric Comprehensive and NLN Maternal-Newborn Comprehensive. Further analysis by logistic regression by Uyehara et al. finding showed the only significant predictor for NCLEX-RN success was the NLN Adult Health Comprehensive Test, $p = .0001$. Thus, results from this investigation and the Uyehara et al. study both show significant positive correlations for Medical-Surgical content assessment testing and NCLEX-RN® outcome.

A weak, direct, positive significant correlation, $r_s(89) = .243$, $p = .021$, was found between Content Mastery Nursing Care of Children (CM-NC) scores and NCLEX-RN ® outcome. This finding indicates that students that scored well on CM-NC assessments were slightly more inclined to be successful on first-attempt NCLEX-RN® testing.

The study by Uyehara et al.(2007) reported significant correlations when comparing NLN Pediatric Comprehensive standardized testing content areas to NCLEX-RN® outcome, though this score was found not significant during logistic regression analysis. Again, though this investigation analyzed ATI rather than NLN assessment products, the studies and this investigation find agreement in significant correlations between pediatric content and NCLEX-RN® outcome.

Between the Content Mastery Mental Health (CM-MHN) and the NCLEX-RN® outcome an extremely weak, direct non-significant correlation that was found. These students' scores would not be useful for the benchmarking process for this program. However, these scores may be helpful for tracking and trending. Findings may suggest that curriculum for this content is in need of evaluation and/or revision. Further research may also call for studies on the ATI CM-MHN test to determine correlation with the NCLEX-RN® test plan. As previously suggested, it is possible that lack of student effort on the Content Mastery assessments may have influenced the correlational values, which may have resulted in the lack of significance for the Mental Health content. Student effort may also have been an influence regarding weaker significant correlations found with other content areas. This may provide guidance for future revision for

benchmarking and/or curriculum design for this program (see Nursing Research and Nursing Education implications discussed later in this chapter).

This investigation was unable to use the Content Mastery Maternal-Newborn (CM-MN) scores due to insufficient number of scores available, thus this analysis was not performed. Therefore, comparison to other studies was not possible.

Ad Hoc Research Question

An additional research inquiry was identified as relevant for this investigation: "Is there a statistically significantly relationship between ATI RN-Predictor Comprehensive sub-scale scores and NCLEX-RN® outcome?" Spearman rho correlation coefficients were calculated for these variables.

Between the scores for RN-Predictor area of Medical-Surgical (P-MS) and NCLEX-RN® outcome, an extremely weak, direct but non-significant correlation was found, $r_s(89) = .073$, $p = .494$ This variable was not found to be a useful tool for benchmarking for this program. However, the data should be tracked and trended to identify patterns. This finding should be investigated further since this investigation found a significant correlation between the two tests over the same medical-surgical (med-surg) content and, in addition, a significant correlation between the exam taken during the curriculum and NCLEX-RN® outcome. Other options should be explored to answer this question. See further discussion under recommendations for future research.

Between the RN-Predictor Mental Health (P-MHN) scores and NCLEX-RN® outcome, a Spearman rho correlation coefficient calculation found an extremely weak, indirect, non-significant correlation, $r_s(89) = -.005$, $p = .966$, indicating that students who

achieve higher scores on the ATI RN-Predictor in the area of Mental Health are more likely to fail on first-attempt NCLEX-RN testing. However, this is impossible to know until further research is conducted. This finding, as with research question one comparing the CM-MHN scores to NCLEX-RN® outcome, indicates that the Mental Health content in this curriculum as well as the evaluation thereof needs further research, tracking and trending. Another potential consideration is the need for further correlational studies between ATI CM-MHN and P-MHN assessments and NCLEX-RN® to determine if these assessments are testing what is intended (see implications for Nursing Research and Nursing Education discussions later in this chapter).

Of additional concern, the NCLEX-RN® 2010 test plan (National Council of State Boards of Nursing [NCBSN], 2010) shows 6 to 12% of the test questions devoted to psychosocial integrity while 39 to 64% of the test is content from physiological integrity. While statistics show that one-fourth of all health disabilities are related to psychiatric or mental health issues or substance abuse, the psychosocial content of entry-level knowledge mastery tested on the NCLEX-RN have decreased by one-half (Poster, 2004). Poster suggested that psychiatric interventions are not reported frequently, perhaps due to the lack of mental health knowledge. Thus, NCSBN views this as information which is not essential to practice. This may result in even less mental health content being taught at nursing programs.

Between the RN-Predictor Maternal-Newborn (P-MN) scores and NCLEX-RN® outcome, a Spearman rho correlation coefficient calculation found a weak, direct

correlation that was not significant, $r_s(89) = .184$, $p = .081$. This ATI testing score is not beneficial for benchmarking use; however, since it is not a significant relationship, the mean scores will provide data for appropriate trending and further investigation into the maternal-newborn content within the curriculum.

A Spearman rho correlation coefficient calculation revealed a weak, direct but significant correlation, $r_s(89) = .207$, $p = .049$ between the RN-Predictor Nursing Care of Children (P-NC) scores and NCLEX-RN® outcome, indicating that students that scored higher in this content area were more likely to succeed on first-attempt NCLEX-RN testing. According to this investigation's findings, this is the only significant RN-Predictor sub-scale score that correlated significantly with NCLEX-RN® outcome, making this a useful tool for benchmarking for this program.

Considering general content areas, the study by Uyehara et al.(2007) revealed significant correlations when comparing NLN Pediatric Comprehensive standardized testing scores to NCLEX-RN® outcome. The corresponding findings indicate that pediatric content assessment scores, taken either at end of didactic learning or at end-of-program, appear to be beneficial for determining NCLEX-RN® outcome.

There is a gap in literature regarding analysis of specific content or sub-scale areas of exit or predictor assessments in correlation NCLEX-RN® outcome, focusing on comprehensive scores instead. Thus, comparison to other studies cannot be done.

<div align="center">Limitations</div>

Due to rural geographical location of the research setting results cannot be generalized to other population settings. This investigation site is a bi-level ADN

program at a community college on two campuses with the same curriculum; results cannot be generalized to other program types and each program must identify its own curriculum needs.

A limitation for this investigation was the delimitation of students who did not complete the program in four consecutive semesters. A pool of 150-200 students was anticipated, but inclusive criteria resulted in total sample of 91. Based on power analysis, to test the research variables, this sample size was sufficient. A larger sample could have provided a larger effect size.

A limitation identified during data collection was that the CM-MN was not utilized consistently throughout the three year time span. With a very low participant number, correlation analyses with P-MN and NCLEX-RN® were not possible.

A limitation discovered during analysis was that the ATI assessment products were updated during the years of data collection (2006-2009). The Content Mastery Series 2.1 was administered for the academic years 2006-2007 and 2007-2008, with the Content Mastery 2007 B starting with students testing 2008-2009. The RN-Predictor version 3.0 was administered for academic year 2006-2007, and the updated version, RN-Predictor 2007, given 2007-2008 and 2008-2009. Due to the sample size, it was not possible to separate the data for correlation by academic years. Therefore, correlations were analyzed for the three academic years combined ($N = 91$). Assessment products have to update assessments with changes based on the evolving NCLEX-RN® exam. This continual change will need to be addressed in future research.

Recommendations

Findings from this investigation indicate the usefulness of the ATI RN-Predictor comprehensive score as a tool for benchmarking, based on its significant correlation with NCLEX-RN® outcome. Both ATI assessments for pediatric content, CM-NC and P-NC, were determined to be significantly correlated with NCLEX-RN® outcome, as well as inter-correlated significantly. This content appears to be directly related to success on first-attempt NCLEX-RN® testing, and useful for benchmarking for this program. No other sub-scale components of the RN-Predictor were found to have significant relationships with NCLEX-RN® outcome, and in fact the P-MHN had an indirect relationship. The CM-MHN also showed non-significance, contributing to the low association of mental health content with NCLEX-RN® outcome, regardless of whether testing occurred during the program or at end-of program. However, analysis of the CM-MHN and P-MHN revealed a low, non-significant relationship, which does not concur with the other findings, as both tests show non-significant (CM-MHN) or even indirect (P-MHN) relationships with NCLEX-RN® outcome. The CM-MS test scores showed a weak but significant relationship with P-MS scores, which is an expected outcome. The CM-MS also had a weak but significant correlation with NCLEX-RN® outcome. Based on these findings, the following recommendations for nursing research, practice, theory and education are presented for consideration.

Nursing Research

Recommendations derived from this investigation include:

1. Continue to maintain an electronic data base of ATI test data for tracking,

    trending and benchmarking.

2. Nursing student exit examinations from most companies break the test into

    various subsections. In addition there was a significant relationship between

    the ATI RN predictor test scores and NCLEX-RN® outcomes so perhaps the

    wrong sub-sections were selected as variables for this investigation. Tracking

    and trending of all the subsections may be able to identify those subsections

    that are more highly correlated with the NCLEX-RN® outcome. For example,

    tracking and trending of specific areas for NCLEX-RN® may be more

    beneficial and the content areas selected for this investigation.

3. Share these current findings with other regional ADN programs for statewide

    networking and curriculum planning.

4. Continue to investigate the relationship between mental health content and

    NCLEX-RN® outcome.

5. Consider further research using discriminant analysis or logistic regression

    analysis on ATI variables to evaluate predictive values for NCLEX-RN®

    outcomes.

Nursing Practice

Because the NCLEX-RN® is used to evaluate mastery of entry-level RN

knowledge and competency, findings from this investigation are useful for guiding this

program to student success on first-time NCLEX-RN® testing. Upon successful completion of nursing program curriculum and licensure, newly graduated ADN students can enter the workforce and utilize the knowledge. Identification of at-risk students and remediation needs during the nursing program can lead to optimal outcomes at end-of program.

Interviewing of selected new graduates may also be helpful in directing curriculum. Novice RNs are closer to their educational experience and may be able to provide appropriate information regarding the needs in the curriculum.

Nursing Theory

This investigation was guided by Imogene King (1971, 1981, 1986), beginning with her Interacting Systems Framework and Theory in Nursing Practice, in which concepts are organized around systems, from personal and interpersonal to larger Social Systems, such as community and educational institutions. This theory would include the personal system of nursing students, interacting through the learning process within the Social Systems, or nursing program; this learning process also includes interpersonal social contact with fellow students and clients. King's (1981) Theory of Goal Attainment speaks to the challenges of student nurses in modern society and the evolving concept of learning. Current society and healthcare needs are continuously changing, and curriculum and learning styles are forced to change as well. Technological advances include computerized assessments of nursing knowledge used by nursing students to progress from potential to reality when goals are met. King (1986) discusses this concept of learning and the need to adapt teaching-learning methods. King's theories and

attributes encompass the nature of the learning process of the nursing student. While this theory supports investigation in nursing education especially goal attainment, King's theory does not suggest specific research questions.

Nursing Education

Results from this investigation will be beneficial for this program of nursing to evaluate and refine curriculum. Data generated supports use of the ATI RN-Predictor comprehensive score as a benchmark for student achievement. Other useful assessment tools identified include the CM-MS assessment and both Nursing Care of Children assessments, CM-NC and P-NC. An identified concern from this data is the low association of Mental Health content with NCLEX-RN® outcome. Recommendations for this nursing program include further investigation into this issue, which will allow curriculum adjustments to align with the NCLEX-RN® test plan.

Continued trending and tracking of ATI assessment data will provide a base of valuable information to guide benchmark goals for significant tests. With benchmark goals set, this will allow the program to consider protocols of progression and completion. Identification of at-risk students during the nursing program will provide sufficient time for individualized remediation plans for assessment improvements before attempting NCLEX-RN® testing. However, the ability to identify such students will require more investigation.

Further research regarding variables which correlate with NCLEX-RN® outcome is encouraged. With increased number of regional ADN program participation, larger sample sizes will provide greater effect with results. Networking between ADN

programs statewide can guide curriculum for the optimal end-result of NCLEX-RN®

success, and therefore, more competent nurses to fill the workforce void.

Summary

Identifying academic and demographic variables that identify at-risk students

early in the nursing education process is essential to successful outcomes. These student

variables include tracked data, such as standardized computerized assessment scores.

With findings of significant correlations with NCLEX-RN® outcome, nursing programs

can begin to set benchmark score goals, individualized plans of study, and consider

progression protocols based on data from research.

REFERENCES

Alexander, J. E., & Brophy, G. H. (1997). A five-year study of graduates' performance

on NCLEX-RN™. *Journal of Nursing Education, 36*(9), 443-445.

Anonymous. (2008). Benchmarking in the health sector. Retrieved on June 4, 2009 at

http://www.aihw.gov.au/publications/hwi/fnrhspi/fnrhspi-c04.pdf

Arathuzik, D., & Aber, C. (1998). Factors associated with National Council Licensure

Examination-Registered Nurse success. *Journal of Professional Nursing, 14*(2),

119-126.

Assessment Technologies Institute ™, LLC. (2005). *RN Content Mastery Series™ 2.1*

Faculty resource information. Overland Park, KS: Assessment Technologies

Institute ™, LLC.

Assessment Technologies Institute ™, LLC. (2007a). *CMS assessment descriptions*.

Overland Park, KS: Assessment Technologies Institute ™, LLC.

Assessment Technologies Institute ™, LLC. (2007b). *Score explanation and*

interpretation of group performance profile. Overland Park, KS: Assessment

Technologies Institute ™, LLC.

Assessment Technologies Institute ™, LLC. (2007c). *Technical manual for the RN*

comprehensive predictor, version 3.0. Overland Park, KS: Assessment

Technologies Institute ™, LLC.

Assessment Technologies Institute ™, LLC. (2007d). *Validation of the RN content*

mastery series™. Overland Park, KS: Assessment Technologies

Institute ™, LLC.

Beeman, P. B., & Waterhouse, J. K. (2001). NCLEX-RN performance: Predicting

    success on the computerized examination. *Journal of Professional Nursing, 17*(4).

    158-165.

Billings, D. M. (2007). Using benchmarking for continuous quality improvement in

    nursing education. *Annual Review of Nursing Education, 5*, 173-180.

Brink, P., & Wood, M. (1994). *Basic steps in planning research: From question to*

    proposal (4th ed.). Boston: Jones & Bartlett.

Briscoe, V. J., & Anema, M. G. (1999). The relationship of academic variables as

    predictors of success on the National Council Licensure Examination for

    Registered Nurses (NCLEX-RN) in a selected associate degree program.

    Association of Black Nursing Faculty Journal, 10(4), 80-84.

Campbell, A. R., & Dickson, C. J. (1996). Predicting student success: A 10-year review

    using integrative review and meta-analysis. *Journal of Professional Nursing,*

    12(1), 47-59.

Carl, L. L. (2007). *Assessment Technologies Institute test scores, NCLEX-RN pass-fail,*

    nursing program evaluation, and catastrophic events in Pennsylvania. (Doctoral

    dissertation, University of Phoenix, 2007). Retrieved on 06/16/10 from

    http://proquest.umi.com.proxy.kumc.edu:2048 (UMI No. 3292150)

Daigle, J. (2007). *Relationship between student NCLEX-RN scores and academic*

    variables. Unpublished master's thesis, Fort Hays State University, Hays, KS.

Daley, L. K., Kirkpatrick, B. L., Frazier, S. K., Chung, M. L., & Moser, D. K. (2003).

 Predictors of NCLEX-RN success in a baccalaureate nursing program as a

 foundation for remediation. *Journal of Nursing Education*, *42*(9), 390.

Davenport, N. C. (2007). A comprehensive approach to NCLEX-RN® success. *Nursing*

 Education Perspectives, 28(1), 30-34.

Fowles, E. R. (1992). Predictors of success on NCLEX-RN and within the nursing

 curriculum: Implications for early intervention. *Journal of Nursing Education,*

 31(2), 53-57.

Giddens, J. F. (2009). Changing paradigms and challenging assumptions: Redefining

 quality and NCLEX-RN pass rates. *Journal of Nursing Education*, *48*(3),

 123-125.

Giddens, J., & Gloeckner, G.W. (2005). The relationship of critical thinking to

 performance on the NCLEX-RN®. *Journal of Nursing Education, 44*(2), 85-90.

Griffiths, M. J., Papastrat, K., Czekanski, K., & Hagan, K. (2004). The lived experience

 of NCLEX failure. *Journal of Nursing Education, 43*(7), 322-325.

Haas, R. A., Nugent, K. E., & Rule, R. A. (2004). The use of discriminant function

 analysis to predict student success on the NCLEX-RN. *Journal of Nursing*

 Education, 43(10), 440-446.

Higgins, B. (2005). Strategies for lowering attrition rates and raising NCLEX-RN® pass

 rates. *Journal of Nursing Education*, *44*(12), 541-548.

Kansas State Board of Nursing. (2008). *Documentation for re-approval of nursing programs in Kansas 60-2-102 through 60-2-107.* Retrieved on 01/08/09 from http://www.ksbn.org/cne/Program%20Re-Approval%20Form.pdf

King, I. M. (1971). *Toward a theory for nursing.* New York: John Wiley.

King, I. M. (1981). *A theory for nursing: Systems, concepts, process.* Albany, NY: Delmar.

King, I. M. (1986). *Curriculum and instruction in nursing: Concepts and process.* Norwalk, CT: Appleton-Century-Crofts.

Kline, T. M. (2010*). Identifying predictors of success for the computerized NCLEX-RN in associate degree nursing students*. (Master's investigational project. Northern Kentucky University, 2010). Retrieved on 06/16/10 from http:/proquest.umi.com.proxy.kumc.edu (UMI No. 1475239)

Morrison, S. (2005). Chapter 5: Improving NCLEX-RN pass rates through internal and external curriculum evaluation. *Annual Review of Nursing Education, 3*, 77-94.

Morrison, S., Free, K. W., & Newman, M. (2002). Do progression and remediation policies improve NCLEX-RN pass rates? *Nurse Educator, 27*(2), 94-96.

Mosser, N. R., Williams, J., & Wood, C. (2006). Use of progression testing throughout nursing programs: How two colleges promote success on the NCLEX-RN®. *Annual Review of Nursing Education, 4*, 305-319.

National Council of State Boards of Nursing, Inc. (2007). *2007 NCLEX-RN® Detailed Test Plan.* Chicago: National Council of State Boards of Nursing, Inc.

National Council of State Boards of Nursing, Inc. (2010). *2010 NCLEX-RN® Detailed Test Plan.* Chicago: National Council of State Boards of Nursing, Inc.

National League for Nursing. (2009). *Excellence Initiatives: Hallmarks, indicators, glossary and references*. Retrieved on 06/04/09 from www.nln.org/excellence/hallmarks_indicators.htm

Newton, S. E., Smith, L. H., Moore, G., & Magnan, M. (2007). Predicting early academic achievement in a baccalaureate nursing program. *Journal of Professional Nursing, 23*(3), 144-149.

Nibert, A., Young, A., & Adamson, C. (2002). Predicting NCLEX success with the HESI Exit Exam: Fourth annual validity study. *Computer Nursing*, *20*(6), 261-267.

Nibert, A. T., Young, A., & Britt, R. (2003). The HESI Exit Exam: Progression benchmark and remediation guide. *Nurse Educator*, *28*(3), 141-145.

Office for Protection from Research Risks. (1997). Summary of basic protections for human subjects. Retrieved December 19, 2007, from www.hhs.gov/ohrp/humansubjects/guidance/backis.htm#Exempt

PASW® for Windows, version 18 by SPSS, (2010). *Statistical Package for Social Sciences*. Chicago, IL:

Polit, D. F., & Hungler, B. P. (1999). *Nursing research, principles, and methods* (6th ed.). Philadelphia: Lippincott.

Poster, E. C. (2004). Psychiatric Nursing at Risk: The New NCLEX-RN Test

Plan. *Journal of Child and Adolescent Psychiatric Nursing*, 17(2), 47-48.

Retrieved June 27, 2010, from Research Library. (Document ID: 678654391).

Rhodes, S. (2006). Relationship between academic performance and NCLEX-RN

performance. Unpublished master's thesis, Fort Hays State University, Hays, KS.

Richards, E. A., & Stone, C. L. (2008). Student evaluation of a standardized

comprehensive testing program. *Nursing Education Perspectives*, *29*(6), 363-365.

Scheckel, M., & Valiga, T. M. (2007). Advancing the science of nursing education:

More findings from the National Survey on Excellence in Nursing Education.

Nursing Education Perspectives, 28(3), 167-170.

Siktberg, L. L., & Dillard, N. L. (2001). Assisting at-risk students in preparing for

NCLEX-RN. *Nurse Educator, 26*(3), 150-152.

Spann, K. (1997). Benchmarking: Best practices. *MedSurg Nursing*. Retrieved on

06/04/2009 from

http://findarticles.com/p/articles/mi_m0FSS/is_n1_v6/ai_n18607252

Spurlock, D. Jr. (2006). Do no harm: Progression policies and high-stakes testing in

nursing education. *Journal of Nursing Education, 45*(8), 297-303.

Spurlock, D. R., & Hanks, C. (2004). Establishing progression policies with the HESI

Exit Examination: A review of the evidence. *Journal of Nursing Education,*

43(12), 539-546.

Spurlock, D. R., & Hunt, L. A. (2008). A study of the usefulness of the HESI Exit Exam

in predicting NCLEX-RN failure. *Journal of Nursing Education*, *47*(4), 157-167.

Uyehara, J., Magnussen, L., Itano, J., & Zhang, S. (2007). Facilitating program and NCLEX-RN success in a generic BSN program. *Nursing Forum*, *42*(1), 31-39.

Washington, L. J., & Perkel, L. (2001). NCLEX-RN strategies for success: A private university's experience. *Association of Black Nursing Faculty Journal, 12*(1), 12-17.

Waterhouse, J. K., & Beeman, P. B. (2003). Predicting NCLEX-RN success: Can it be simplified? *Nursing Education Perspectives, 24*(1), 35-40.

Wong, J., & Wong, S. (1999). Contribution of basic sciences to academic success in nursing education. *International Journal of Nursing Studies, 36*, 345-354.

Yin, T., & Burger, C. (2003). Predictors of NCLEX-RN success of Associate Degree Nursing graduates. *Nurse Educator, 28*(5), 232-236.

Zuzelo, P. (1999). Professional practice and the NCLEX examination: A bottom line approach. *Nurse Educator, 24*(3), 11-12.

Appendix A:

NREC Approval for Investigation

Date: 12/11/2009 11:47 AM
To: "Karolyn Kells" <kkells@fhsu.edu>, "Karen Trible" <ktrible@fhsu.edu>, "Amy Waters" <ajwaters@scatcat.fhsu.edu>
From: "Leslie Paige" <no-reply@irbnet.org>
Reply To: "Leslie Paige" <lpaige@fhsu.edu>
Subject: IRBNet Board Action

Please note that Fort Hays State University IRB has taken the following action on IRBNet:

Submission: [122803-1] RELATIONSHIPS BETWEEN NURSING EDUCATION VARIABLES
Action: EXEMPT
Effective Date: November 18, 2009

Should you have any questions you may contact Leslie Paige at lpaige@fhsu.edu.

Appendix B:

Approval to Collect Data

June 8, 2009
Pratt Community College Chandler School of Nursing

Gail Withers, Dean of Nursing,

For my final Master's degree research project, I am proposing an investigation to correlate NCLEX-RN outcomes with student scores on the Assessment Technologies Institute ™ (ATI) Content Mastery Assessment scores and the ATI RN Predictor scores.

I am asking for rights and privileges from your nursing program and college administration to collect student data from the academic years of 2006-2007, 2007-2008, and 2008-2009. Data of interest for this investigation include archived ATI scores and NCLEX-RN results for these graduated students.

I will be using a research assistant, who is a Pratt Community College Chandler School of Nursing faculty member, to code student names to maintain confidentiality and apply this data to worksheets for my use in statistical analysis. Student data will not be connected to names or other identifiers at any point after coding. There are no anticipated risks or monetary costs to your program for this investigation.

This data collection time period is expected to take approximately 10-14 working days. Data will be stored and protected under lock and key. Abstracted data from your program will be used for research purposes and findings will appear in my published thesis report and/or other journal publication. However, all data will be report in aggregate form. Your school and program name will not be used in future publications.

At the conclusion of this investigation, copies of all data sets will be provided to the research setting at the Chandler School of Nursing. Findings will be presented to your faculty.

Benefits of this investigation for your nursing program include important information for your faculty to use for curriculum improvements and benefits for future students and NCLEX-RN success. Findings may also be useful for your department to use in your next accreditation visit process.

For any questions now or during this investigation, please feel free to contact me via phone or e-mail, or my thesis chairperson at Fort Hays State University, Dr. Karolyn Kells at kkells@fhsu.edu, or phone 785-628-4519.

Please check your consent selection below and return to me with your dated signatures.

Thank you,
Amy J. Waters
amy.waters@gcccks.edu
620-770-0479

☒ Pratt Community College approves this investigation as described above.

☐ Pratt Community College does not approve this investigation.

_____   Date 11/11/09 _____

Gail L. Withers _____   Date 11/11/09 _____

Appendix C:

Data Collection Worksheet

DATA COLLECTION WORKSHEET

Coded Student ID Number _____

ADN graduation year 20_____

ATI CONTENT MASTERY SCORES

_____ Medical-Surgical Nursing

_____ Maternal-Newborn Nursing

_____ Nursing Care of Children

_____ Mental Health Nursing

ATI RN-PREDICTOR SCORE _____

PREDICTOR SUB-SCALE SCORES

_____ Medical-Surgical

_____ Maternal-Newborn Nursing

_____ Nursing Care of Children

_____ Mental Health Nursing

NCLEX-RN OUTCOME _____PASS _____FAIL

AGE _____

GENDER _____

ETHNICITY _____