HEALTH NUMERACY AND COMMUNICATING RISK TO PATIENTS:
MEDICAL STUDENTS’ RISK LITERACY AND ATTITUDES TOWARD STATISTICS

by

Tina Dalene Moore
Liberty University

A Dissertation Presented in Partial Fulfillment
Of the Requirements for the Degree
Doctor of Education

Liberty University
2017
HEALTH NUMERACY AND COMMUNICATING RISK TO PATIENTS:
MEDICAL STUDENTS’ RISK LITERACY AND ATTITUDES TOWARD STATISTICS

by Tina Dalene Moore

A Dissertation Presented in Partial Fulfillment
Of the Requirements for the Degree
Doctor of Education

Liberty University, Lynchburg, VA
2017

APPROVED BY:

Michelle Barthlow, Ed.D., Committee Chair

Sarah Horne, Ed.D., Committee Member

Kristie Hadden, Ph.D., Committee Member
ABSTRACT

Numeracy, a domain of health literacy, poses barriers to effective patient-provider communication including understanding risks and benefits for shared decision making. Best practices for conveying numeric information have been established; however, these skills are not embedded in current medical school curricula. Inadequate statistical literacy of physicians can lead to misinterpretation of risk and benefit data as well as poor communication with patients. Medical school curricula and pre-matriculation requirements for statistics varies. This quantitative, non-experimental, predictive, correlational study explored medical students’ risk literacy and attitudes toward statistics. The purpose was to examine the relationship between risk literacy as measured by the Berlin Numeracy Test and a linear combination of the Attitudes Toward Statistics (ATS) subscales and successful completion of a college level statistics course. The sample included 327 first year medical students from two U.S. academic institutions located in different states. Survey data were entered into SPSS; initial data screening and assumption tests were conducted prior to multiple regression analysis. The linear combination of predictor variables had a significant yet modest correlation with the criterion variable. Both attitude subscales had a statistically significant positive predictive relationship with risk literacy, while the impact of a college level statistics course was insignificant; over 90% of the sample had completed a college level statistics course which may have led to inconclusive findings for that variable. Participants demonstrated a lack of risk literacy; only 34.9% answered more than half correctly. Recommendations for future research are included and should primarily focus on developing and testing risk literacy and risk communication training.

Keywords: health literacy, health numeracy, medical students, shared decision-making, risk literacy, statistics
Dedication

This dissertation is dedicated to my family, those in my home supporting and encouraging me through this doctoral journey and my parents who instilled a strong work ethic in me from the beginning. The sacrifices made along the way were certainly not a burden carried alone, as my husband and children also were impacted by my pursuit of this degree. More nights and weekends than I want to remember were spent with my laptop as my closest companion rather than with those who needed me the most. They never complained but always tried to be understanding of the demands of my program. With the time requirements of a full-time job and school, my family only got the fumes of me that remained. I know the cost to my family was much more than just money, as the price paid included family time and often mom’s sanity! Thank you Chad, Blakely, and Keifer for understanding my desire to achieve this goal in life and never complaining about what it took away from you. Be ready to call me Dr. Mom for at least a while.

I started this journey with two, living parents but am ending with both having gone to be with the Lord. Still yet, they continued to inspire me to complete my degree and reach my fullest potential. Neither of my parents was well educated, and, in fact, neither completed high school. My dad worked hard at a blue collar job throughout my childhood to provide for us, and my mom budgeted our money carefully so we could live well on a small income. I truly can’t remember my dad ever taking a sick day to miss work. He was a man of few words but a man of his word. My mom spent her entire life doing for others; she had a true servant heart. The complement of my dad, she was a woman of many words! Together they raised me in a strong, Christian home, and their values made me the person I am proud to be today. Thank you mom and dad!
Acknowledgments

Little did I know when I started my doctoral program how I would end it, but God has plans beyond our imagination that work for our good. I began thinking my dissertation would be related to secondary math education; however, my career took a different route once I reached the dissertation phase of the program, and I was blessed to be able to explore a research topic that coincided with new role in health literacy. I would like to thank Dr. Kristie Hadden for giving me opportunities at the University of Arkansas for Medical Sciences Center for Health Literacy (UAMS CHL) that have enabled me to train healthcare professionals and learners and focus my efforts on medical school curricula, numeracy, and shared decision making. I can’t begin to fathom what lies ahead for UAMS CHL and our work in these areas of research.

I would also like to thank Dr. Cliff Coleman for his partnership that enabled me to collect data from University B, without which, I could not have conducted this research. He is a national thought leader on health literacy and medical school curriculum, yet he took the time to assist me with this study. I hope the analysis of this data will result in valuable information to guide curricular decisions at his institution and University A as well as other colleges of medicine.

Additionally, I am thankful for my chair, Dr. Michelle Barthlow, and committee members, Dr. Sarah Horne and Dr. Hadden, who provided prompt, relevant, and helpful feedback to improve my work. These brilliant women were not only pillars of wisdom due to their academic and content expertise, but they also saved me from the horror stories so many doctoral candidates report. Dr. Barthlow was a spiritual mentor, offering support and encouragement since the inception of my research question during EDUC 980. I could not have asked for more from a chair than “Dr. B” has given of herself during my journey as a doctoral candidate.
# Table of Contents

ABSTRACT .....................................................................................................................................3

Dedication ........................................................................................................................................4

Acknowledgments ............................................................................................................................5

List of Tables ...................................................................................................................................9

List of Figures ................................................................................................................................10

List of Abbreviations .....................................................................................................................12

CHAPTER ONE: INTRODUCTION ............................................................................................13

  Overview ...................................................................................................................................13

  Background .................................................................................................................................13

  Problem Statement .......................................................................................................................19

  Purpose Statement .......................................................................................................................21

  Significance of the Study ............................................................................................................22

  Research Question(s) ..................................................................................................................23

  Definitions .................................................................................................................................24

CHAPTER TWO: LITERATURE REVIEW ................................................................................26

  Overview ...................................................................................................................................26

  Theoretical Framework ...............................................................................................................26

  Related Literature .......................................................................................................................28

  Summary ...................................................................................................................................59

CHAPTER THREE: METHODS ..................................................................................................60

  Overview ...................................................................................................................................60

  Design .......................................................................................................................................60
APPENDIX B: Permission to Use Berlin Numeracy Test .............................................126
APPENDIX C: Attitudes Toward Statistics Scale..........................................................127
APPENDIX D: Permission to Use Attitudes Toward Statistics scale ..............................129
APPENDIX E: Email Request to Include University A Students in Research .................130
APPENDIX F: Email Request to Include University B Students in Research...............131
List of Tables

Table 1: Statistics Course Admission Prerequisite Requirements of 27 U.S. Medical Schools...58
Table 2: Means and Standard Deviations of Variables by University and Overall..................73
Table 3: Means and Standard Deviations of Variables by Demographic...............................74
Table 4: Attitude Toward the Field of Statistics Subscale Response Percentages.......................77
Table 5: Attitude Toward the Course of Statistics Subscale Response Percentages....................79
Table 6: Pearson Correlation Coefficients of Independent Variables.................................87
Table 7: Variance Inflation Factors......................................................................................87
List of Figures

Figure 1: Percentage of U.S. adults in lowest levels of numeracy and literacy on the NAAL and the PIAAC…………………………………………………………………..……….16

Figure 2: Health literacy levels of U.S. adults on the NAAL…………………………...……17

Figure 3: Percentage of U.S. adults expressing lack of confidence in understanding medical statistics by race on the HINTS………………………………………………………………….18

Figure 4: Percentage correct and incorrect by question on the Berlin Numeracy Test………..76

Figure 5: Percentages of scores on the Berlin Numeracy Test………………………………………76

Figure 6: Responses to the question, “Did you successfully complete a college level statistics course with a grade of C or better?”…………………………………………………………………………………..80

Figure 7: Histogram of Berlin Numeracy Test scores………………………………………………82

Figure 8: Histogram of Attitude Toward the Field of Statistics scores………………………….82

Figure 9: Histogram of Attitude Toward the Course of Statistics scores……………………….83

Figure 10: Scatterplot of Attitude Toward the Field of Statistics and Berlin Numeracy Test scores………………………………………………………………………………….84

Figure 11: Scatterplot of Attitude Toward the Course of Statistics and Berlin Numeracy Test scores………………………………………………………………………………….84

Figure 12: Plots of the standardized residuals and the predicted standardized residuals for the regression model with the predictor variable Attitude Toward the Field of Statistics and criterion variable Berlin Numeracy Test………………………………………………………….85

Figure 13: Plots of the standardized residuals and the predicted standardized residuals for the regression model with the predictor variable Attitude Toward the Course of Statistics and criterion variable Berlin Numeracy Test………………………………………………………….85
Figure 14: Plots of the standardized residuals and the predicted standardized residuals for
the regression model with the predictor variable Successful Completion of a College Level
Statistics Course and criterion variable Berlin Numeracy Test.................................86
Figure 15: Normal probability plot of regression standardized residuals for Hypothesis 1……88
Figure 16: Normal probability plot of regression standardized residuals for Hypothesis 2……88
Figure 17: Normal probability plot of regression standardized residuals for Hypothesis 3……89
Figure 18: Normal probability plot of regression standardized residuals for Hypothesis 4……89
Figure 19: Histogram of regression standardized residuals for Hypothesis 1.......................90
Figure 20: Histogram of regression standardized residuals for Hypothesis 2.......................90
Figure 21: Histogram of regression standardized residuals for Hypothesis 3.......................91
Figure 22: Histogram of regression standardized residuals for Hypothesis 4 .......................91
List of Abbreviations

Agency for Healthcare Research and Quality (AHRQ)
American Association of Medical Colleges (AAMC)
American Board of Internal Medicine (ABIM)
Attitudes Toward Statistics (ATS)
Health Information National Trends Survey (HINTS)
Institute of Medicine (IOM)
Institutional Review Board (IRB)
Liaison Committee on Medical Education (LCME)
National Assessment of Adult Literacy (NAAL)
Organization for Economic Cooperation and Development (OECD)
Program for the International Assessment of Adult Competencies (PIAAC)
Statistics Attitude Survey (SAS)
Survey of Attitudes Toward Statistics (SATS)
United States Medical Licensing Examination (USMLE)
U.S. Food and Drug Administration (FDA)
CHAPTER ONE: INTRODUCTION

Overview

This chapter introduces the role of numeracy in health with a focus on patient-provider communication for informed decision making and statistical literacy of medical professionals. The background of the topic, the problem statement, the purpose of the study, and the significance of the study are presented. In addition, the research question and hypotheses, definitions of terms, assumptions, and limitations are discussed.

Background

Only 12% of the American population has proficient health literacy (Kutner, Greenberg, Jin, & Paulsen, 2006), and individuals lack quantitative literacy skills even more than prose literacy (Goodman, Finnegan, Mohadjer, Krenzke, & Hogan, 2013; Kutner et al., 2006). Lower health literacy and lower health numeracy are associated with poorer health status and negative health outcomes (Berkman et al., 2011; Garcia-Retamero, Andrade, Sharit, & Ruiz, 2015; Omachi, Sarkar, Yelin, Blanc, & Katz, 2013). Health numeracy, a component of overall health literacy, is “the degree to which individuals have the capacity to access, process, interpret, communicate and act on numerical, quantitative, graphical, biostatistical, and probabilistic health information needed to make effective health decisions” (Golbeck, Ahlers-Schmidt, Paschal, & Dismuke, 2005, p. 375). Throughout health care, the need for patients to use, understand, and work with numbers is evident. Examples include choosing a health insurance plan, following medication dosing instructions, self-managing chronic diseases, and understanding risks and benefits for informed decision-making (Garcia-Retamero & Cokely, 2013; Garrison et al., 2012; Golbeck et al., 2005; French, 2014; Huang, Chan, & Feng, 2012; Zikmund-Fisher, Exe, & Witteman; 2014).
Evidence-based best practices for communicating numeric information to patients, including risks and benefits, have been suggested in peer-reviewed literature (Ahmed, Naik, Willoughby, & Edwards, 2012; Fagerlin, Zikmund-Fisher, & Ubel, 2011; Fischhoff, Brewer, & Downs, 2011; Gaissmaier et al., 2012). However, studies have indicated health professionals often do not understand health statistics themselves (Anderson, Gigerenzer, Parker, & Schulkin, 2014; Bookstaver et al., 2012; Gigerenzer & Gray, 2011; Susarla & Redett, 2014; Wegwarth, Schwartz, Woloshin, Gaissmaier, & Gigerenzer, 2012). The statistical numeracy of health professionals impacts how they interpret medical data and how they communicate numeric risk and benefit information to patients (Anderson, Obrecht, Chapman, Driscoll, & Schulkin, 2011; Anderson & Schulkin, 2011; Garcia-Retamero & Hoffrage, 2013; Wegwarth, Gaissmaier, & Gigerenzer, 2011; Wegwarth et al., 2012).

Addressing health literacy and improving patient-provider communication for informed decision making and self-management are priority areas for improving health care quality (Adams & Corrigan, 2003; U.S. Department of Health and Human Services, 2010a). Some research has explored and found need for improvement of both the knowledge and use of health literacy communication best practices of medical professionals (Ali, Ferguson, Mitha, & Hanlon, 2014; Cafiero, 2013; Cailor & Chen, 2015; Howard, Jacobson, & Kripalani, 2013). Health professionals often overestimate their ability to effectively communicate with patients (Howard et al., 2013). Other research has examined the current training curricula of medical professionals and found the focus on health literacy competencies generally lacking (Coleman, 2011; Coleman & Appy, 2012; Henry, Holmboe, & Frankel, 2013). Recent studies have suggested training interventions and competencies for medical school curricula that focus on clear communication and health literacy best practices, but little attention is given explicitly to
Building upon current research that is serving as a catalyst to embed health literacy into medical professions curricula, additional research is needed to examine patient-provider communication related to numeracy, specifically risk communication (Brust-Renck, Royer, & Reyna, 2013; Neuner-Jehle, Senn, Wegwarth, Rosemann, & Steurer, 2011).

A first step in medical school curricular change is determining where gaps exist. This includes problem identification, general needs assessment, and needs assessment of the learners (Thomas, Kern, Hughes, & Chen, 2015). The problem of health numeracy as a barrier to effective patient-provider communication and patient understanding is well established (Brust-Renck et al., 2013; Zipkin et al., 2014). The majority of patients do not have proficient health numeracy, and physicians should use evidence-based best practices to communicate numeric information with all patients but often don’t (Ahmed et al., 2012; Fagerlin et al., 2011; Fischhoff et al., 2011; Gaismaier et al., 2012). Therefore, a general need exists to educate medical students on using those best practices. The literature suggests a gap in the curriculum for teaching medical professionals how to effectively communicate risk (Han et al., 2014), but a possible underlying issue of low statistical literacy of medical professionals is overlooked. Teaching medical professionals how to communicate risk to patients may be ineffective if the medical professionals do not understand the statistics and probabilities of risk.

The 2003 National Assessment of Adult Literacy (NAAL) was administered to over 19,000 adults age 16 and older in the U.S. and found that 22% of adults have below basic quantitative skills while another 33% have only basic quantitative skills (Kutner et al., 2006). The Program for the International Assessment of Adult Competencies (PIAAC) through the
Organization for Economic Cooperation and Development (OECD) was administered to 724 million adults age 16 to 65 in 23 countries including 5,000 adults in the U.S. and found that Americans scored higher in numeracy than 2 countries, lower in numeracy than 18 countries, and statistically the same as 2 countries (Goodman et al., 2013). Participants performed significantly lower on quantitative tasks than prose literacy tasks on both the NAAL (Kutner et al., 2006) and the PIAAC (Goodman et al., 2013) (see Figure 1).

**Figure 1.** Percentage of U.S. adults in lowest levels of numeracy and literacy on the NAAL and the PIAAC. NAAL data represents Below Basic percentages for quantitative versus prose literacy categories. PIAAC data represents Level 1 and below percentages for numeracy versus literacy categories. (Goodman et al., 2013; Kutner et al., 2006)

The 2003 NAAL was the first national assessment that also specifically measured health literacy (Kutner et al., 2006). The health literacy portion of the NAAL included 12 prose literacy
items, 12 document literacy items, and 4 quantitative literacy items. Difficulty level and score reports fell into four categories: below basic, basic, intermediate, and proficient. For example, reading a BMI (body mass index) chart was considered an intermediate skill while circling the date of an appointment on a medical slip was considered a below basic skill. Thirty-six percent of participants fell into the basic or below basic health literacy categories (see Figure 2).

![Health literacy levels of U.S. adults on the NAAL. (Kutner et al., 2006)](chart)

Figure 2. Health literacy levels of U.S. adults on the NAAL. (Kutner et al., 2006)

The 2007 Health Information National Trends Survey (HINTS) further examined aspects of health literacy such as how people access and use health information, manage health, and engage in healthy behaviors (Smith, Wolf, & Wagner, 2010). The cross-sectional survey of American adults included questions such as, “In general, how easy or hard do you find it to understand medical statistics?” on a four-point Likert scale with the choices very easy, easy, hard, or very hard (Huang et al., 2012, p. 5). Analysis of HINTS data revealed 37.4% of individuals lacked confidence in understanding medical statistics with Hispanics and blacks
reporting less confidence than non-Hispanic whites (Huang et al., 2012; Smith et al., 2010) (see Figure 3).

Figure 3. Percentage of U.S. adults expressing lack of confidence in understanding medical statistics by race on the HINTS. (Huang et al., 2012)

The 2010 National Action Plan to Improve Health Literacy (U.S. Department of Health and Human Services, 2010b) and the 2004 Institute of Medicine (IOM) report, Health Literacy: A Prescription to End Confusion (Nielsen-Bohlman, Panzer, & Kindig, 2004) called for health professions education programs to include training in health literacy as part of the curriculum. Healthy People 2020, another federal initiative in 2010, included communication objectives to increase health literacy and promote shared decision making (U.S. Department of Health and Human Services, 2010a). Low health numeracy can be a barrier to shared decision making when patients do not understand numerical concepts of risks and benefits (Elwyn et al., 2012). RTI
International- University of North Carolina Evidence-based Practice Center commissioned by the Agency for Healthcare Research and Quality (AHRQ) updated the 2004 systematic review of health literacy interventions and outcomes (Berkman et al., 2011) and included interventions for numeracy such as best practices for communicating risk to patients. However, research has shown that physicians are not being adequately trained in this area (Han et al., 2014) or in health literacy best practices in general (Ali, 2012; Coleman & Appy, 2012; Toronto & Weatherford, 2015).

In addition to laypeople’s lack of numeracy, since the 1980’s, studies have shown physicians lack the biostatistical knowledge needed for interpreting results from clinical studies (Anderson et al., 2014; Berwick, Fineberg, & Weinstein, 1981; Weiss & Samet, 1980; Windish, Huot, & Green, 2007; Wulff, Anderson, Brandenhoff, & Guttler, 1987). Despite identifying the problem over three decades ago, the deficiency continues to be a problem (Susarla & Redett, 2014). Studies examining physician numeracy have primarily assessed biostatistical rather than general statistical numeracy and risk literacy. Although biostatistical knowledge is needed to interpret results from clinical studies, physicians must have a solid grasp of basic statistics to understand and communicate risk to patients (Caverly et al., 2015). Most U.S. medical schools do not require a statistics course for pre-matriculation (see Table 1), and most medical school curricula do not include instruction in basic statistics (Wegwarth & Gigerenzer, 2011).

Problem Statement

Research has examined medical professional curricula and found a gap in health literacy communication competencies (Coleman, 2011; Coleman & Appy, 2012; Henry, Holmboe, & Frankel, 2013). While health literacy training interventions have been developed for medical professionals, most do not specifically address numeracy or risk communication (Coleman &
Fromer, 2015; Coleman et al., 2013; Green et al., 2014; Henry et al., 2013). The literature does not thoroughly review medical professional curricula for statistics and risk competencies yet has established the need to improve communication of risk to patients (Berkman et al., 2011; Han et al., 2014).

Physicians need strong quantitative literacy, specifically, statistical literacy, to effectively communicate risk and benefit information to patients for informed decision making (Ahmed et al., 2012; Caverly et al., 2015; Han et al., 2014; Neuner-Jehle et al., 2011). Physicians’ lack of numeracy skills could contribute to poor communication of risks and benefits with patients (Anderson & Schulkin, 2011; Garcia-Retamero & Hoffrage, 2013). Probability and statistics are a higher-level domain within numeracy and are more challenging than general numeracy (Golbeck et al, 2005). Most U.S. medical schools do not require college level statistics courses as a pre-matriculation requirement (see Table 1). Research has examined biostatistical knowledge of physicians but has not provided as much focus on the prerequisite of understanding risk and general statistics (Anderson et al., 2014; Berwick et al., 1981; Susarla & Redett, 2014; Weiss & Samet, 1980; Windish, Huot, & Green, 2007; Wulff et al., 1987). Studies suggest that physicians are completing medical school without a solid understanding of biostatistics (Susarla & Redett, 2014), but the literature has not established where in the medical school curriculum students are expected to gain risk literacy. Typically, coursework includes biostatistics, which requires a foundational knowledge of general statistics (Wegwarth & Gigerenzer, 2011).

Students in health professions training, including medical school, often have anxiety and negative attitudes toward statistics (Beurze, Donders, Zielhuis, de Vegt, & Verbeek, 2013; Bookstaver et al., 2012; Hannigan, Hegarty, & McGrath, 2014; Kiekkas et al., Bookstaver et al.,
2012; Johnson et al., 2014; Kiekkas et al., 2015; Zhang, 2012), yet the curriculum does not appear to address the issue. To effectively communicate risk to patients, healthcare professionals must first have a solid understanding of risk themselves. A strong foundational knowledge of and positive attitudes toward statistics are needed to learn biostatistics and the fundamentals of evidence based medicine. The problem is that patients lack health numeracy skills needed to understand risk and physicians aren’t adequately trained in medical school to help them make important health decisions based on risk. With a better understanding of the barriers physicians face in communicating risk effectively, like their competency in statistics, medical schools can address the need for better training that will likely have an impact on improved health for patients. More research is needed to explore the relationship between medical students’ risk literacy and attitudes toward statistics and determine if a gap in the curriculum exists for this competency (Anderson et al., 2011; Levy, Ubel, Dillard, Weir, & Fagerlin, 2013).

**Purpose Statement**

The purpose of this quantitative, non-experimental, predictive, correlational study is to examine the relationship between medical students’ risk literacy as measured by the Berlin Numeracy Test and a linear combination of The Attitudes Toward Statistics (ATS) subscales and successful completion of a college level statistics course (Warner, 2013). The three predictor variables are score on the Attitude Toward the Field of Statistics subscale, score on the Attitude Toward the Course of Statistics subscale, and successful completion of a college level statistics course. The criterion variable is score on the Berlin Numeracy Test. The ATS subscales measure “two distinct aspects of student attitudes toward statistics” (Wise, 1985, p. 404). The Field subscale measures the attitude toward the usefulness of statistics either in general or specifically in the student’s field of study. The Course subscale measures the attitude toward a
Statistics course. Successful completion of a college level statistics course is a self-reported measure, yes or no, to completing a college level statistics course with a grade of C or better. The Berlin Numeracy Test measures statistical numeracy and risk literacy in educated and highly-educated samples (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012). Statistical numeracy is “an understanding of the operations of probabilistic and statistical computation, such as comparing and transforming probabilities and proportions” (Cokely et al., 2012, p. 25). Risk literacy is the “ability to accurately interpret and act on information about risk” (Cokely et al., 2012, p. 26). The population was first year medical students, and the sample was from two U.S. academic institutions located in different states. Participants self-reported as 44.0% male, 56.0% female, and 0% other gender; 7.1% Hispanic/ Latino/a, or Spanish origin and 92.9% not; and 3.1% Black/African American, 0.9% American Indian/Native Alaskan, 0.9% Native Hawaiian/Pacific Islander, 81.7% White, 13.0% Asian, and 4.3% other race.

**Significance of the Study**

Improving patient-provider communication, addressing health literacy and numeracy, and promoting shared decision making are all current goals for health care (Koh et al., 2012; Nielsen-Bohlman et al., 2004; U.S. Department of Health and Human Services, 2010a; U.S. Department of Health and Human Services, 2010b). Yet research suggests current curricula are not preparing physicians to use health literacy best practices nor to communicate risk (Ali et al., 2014; Coleman & Appy, 2012; Han et al., 2014; Karsenty, Landau, & Ferguson, 2013). More research is needed regarding teaching interventions for health literacy and communication with patients (Coleman & Fromer, 2015; Toronto & Weatherford, 2015). Limited curricula have been developed for health literacy interventions and even less for communication of risk (Cailor & Chen, 2015; Coleman & Fromer, 2015; Green et al., 2014; Han et al., 2014).
Research suggests that physicians lack biostatistical knowledge, but few studies have addressed physicians’ general statistical literacy (Anderson et al., 2014; Berwick et al., 1981; Weiss & Samet, 1980; Windish et al., 2007; Wulff et al., 1987). Research has not established if general statistical literacy is acquired along the way through medical school or is static based on prerequisite knowledge. It is premature to develop training on communicating risk prior to determining the level of medical students’ statistical literacy. Some research has examined health professions students’ knowledge of and attitudes toward statistics but not specifically risk literacy (Beurze et al., 2013; Bookstaver et al., 2012; Hannigan et al., 2014; Johnson et al., 2014; Kiekkas et al., 2015; Zhang, 2012). Examining the correlation of attitudes toward statistics and successful completion of a college level statistics course with medical students’ risk literacy is a first step in curriculum needs assessment in order to achieve the goals of improving patient-provider communication, addressing health literacy and numeracy, and promoting shared decision making (Han et al., 2014; Thomas et al., 2015). Determining if medical students have negative attitudes toward statistics and if those attitudes are related to prerequisite training could guide curriculum decisions such as whether to require a college level statistics course for prematriculation. Furthermore, if the study reveals medical students have a deficit in statistical literacy, this could justify the need for curricular change to include either prematriculation requirements for a college level statistics course or medical school instruction in basic statistics prior to biostatistics coursework and topics of evidenced-based medicine.

**Research Question(s)**

**RQ1:** How accurately can risk literacy as measured by the Berlin Numeracy Test be predicted from a linear combination of score on the Attitude Toward the Field of Statistics
subscale, score on the Attitude Toward the Course of Statistics subscale, and successful completion of a college level statistics course?

**Definitions**

1. Biostatistical literacy- understanding of the “treatment and analysis of numerical data derived from biological, biomedical, and health-related studies…It is a way to detect patterns and judge responses.” (Gerstman, 2014, p. 3)

2. Health literacy- “the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions” (Nielsen-Bohlman et al., 2004, p. 32)

3. Health numeracy- “the degree to which individuals have the capacity to access, process, interpret, communicate, and act on numerical, quantitative, graphical, biostatistical, and probabilistic health information needed to make effective health decisions” (Golbeck et al., 2005, p. 375)

4. Risk literacy- “the ability to accurately interpret and act on information about risk” (Cokely et al., 2012, p. 26)

5. Shared decision making- “the process through which clinicians and patients share information with each other and work toward decisions about treatment chosen from medically reasonable options that are aligned with the patients’ values, goals, and preferences” (Allen et al., 2012, p. 2)

6. Statistical literacy- “understanding the statistical aspects and terminology associated with the design, analysis, and conclusions of original research” (Anderson, Williams, & Schulkin, 2013, p. 272)
7. Statistical numeracy- “understanding of the operations of probabilistic and statistical computation, such as comparing and transforming probabilities and proportions” (Cokely et al., 2012)

8. Teach-back- “… a way of checking understanding by asking patients to state in their own words what they need to know or do about their health … a way to confirm that you have explained things in a manner your patients understand” (Brega et al., 2015, p. 18)
CHAPTER TWO: LITERATURE REVIEW

Overview

This chapter starts with an explanation of the theoretical framework that guided this study followed by a synthesis of the related literature that justifies the significance of this research and ends with a summary of the chapter. The literature review provides an appraisal of health literacy and numeracy, and evidence-based best practices for communicating risk are outlined. Health professional knowledge and use of health literacy and numeracy best practices are reported along with an evaluation of the current curriculum on these skills. Shared decision making is explained and linked to health literacy and numeracy. The statistical numeracy of both health professionals and health professions students is reviewed. Finally, medical school curriculum for statistics is explored including data of prerequisite requirements for 27 U.S. academic institutions.

Theoretical Framework

The theoretical framework to guide this study is based upon Robert Gagne’s learning theory that ascertains the mastery of higher-level skills requires prior mastery of lower-level skills (Gagne & White, 1978). Gagne explained learning as a three-part process starting with instruction, development of a memory structure, and ending with the learning outcome. He classified two categories of learning outcomes, retention and transfer of learning, and identified five domains of learning outcomes: information, intellectual skills, cognitive strategies, motor skills, and attitudes. According to Gagne, learning skills in one domain may not necessarily impact the learning of skills in another domain.

The hierarchal information processing suggested by Gagne supports this research because learning to communicate risk to patients is an advanced learning outcome with numerous
prerequisite skills that must be mastered first (Gagne & White, 1987). Prior to learning how to communicate risk, one must have a strong foundation in understanding risk and statistics, especially those found in biostatistics (Caverly et al., 2015). Course work in biostatistics assumes prior knowledge of basic statistics (Trickey, Crosby, Singh, & Dort, 2014). Effective communication of risk to patients is not possible without adequate statistical literacy of the medical professional. However, there appears to be a gap in medical school curricula due to lack of vertical alignment of prerequisite statistical skills, coursework in medical school on statistics and/or biostatistics, and coursework on how to effectively communicate risk to patients. The latter skill depends on mastery of the aforementioned skills.

Including the predictor variables on attitudes toward statistics in this study aligns with one of Gagne’s domains of learning outcomes. Gagne proposed one learning domain may not necessarily influence another; therefore, the ATS scale was analyzed by subscales rather than in its entirety. For example, a medical student could complete a statistics course, gain the intellectual skills needed, and have a positive attitude toward the course of statistics but still have a negative attitude toward the field of statistics. Wise (1985) did not include a discussion of a theoretical framework on which the ATS scale was constructed.

Cokely et al. (2012, p. 37) discussed the causal frameworks “between numeracy, risk literacy, and risky decision making” considered during the development of the Berlin Numeracy Test. In line with Gagne’s theory, Cokely et al. (2012, p. 38) stated “causal differences in risky decisions result from independent contributions of both evolved and acquired numerical estimation systems.” Intuitive number sense and the affective meaning from it can differ among individuals and can vary by situation.
Additionally, Cokely et al. (2012) referred to Reyna and colleagues’ theoretical framework, fuzzy trace theory, a dual-processing model that differentiates between analytical processing of verbatim (literal, factual) information and gist (imprecise, general) information (Reyna, Nelson, Han, & Dieckmann, 2009). Although simultaneous encoding occurs, people tend to have a preference for gist information processing, referred to as fuzzy processing. This type of information processing can influence judgment and decision making.

Finally, developers of the Berlin Numeracy Test described a third framework to understand numeracy involving information processing theory regarding computational approaches. From this perspective, cognitive process tracing studies such as “reaction times, eye-tracking, information search, and think-aloud protocols” are often used to determine “how cognition unfolds over time” shedding light on strategies and mechanisms used (Cokely et al., 2012, p. 39). Following discussion of the three causal frameworks, the test developers summarize that the complex interaction of many factors impacts individual risk literacy and risky decision making.

**Related Literature**

**Health Literacy and Numeracy**

**The impact of low health numeracy of patients.** The results from the 2003 NAAL (Kutner et al., 2006), the 2007 HINTS (Smith et al., 2010), and the PIAAC (Goodman et al., 2013) revealed the dismal status of health literacy and numeracy in the U.S. Health literacy is “the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions” (Nielsen-Bohlman et al., 2004, p. 32). Over 80 million Americans have inadequate health literacy skills to navigate the complex demands of the health care system (Kutner et al., 2006). Health numeracy refers to
components of health literacy that require quantitative skills such as selecting an insurance plan based on copays and deductibles, dosing medication, counting carbohydrates for glycemic control in diabetes, and understanding risk for informed decision making (Garcia-Retamero & Cokely, 2013; Garrison et al., 2012; Golbeck et al., 2005; French, 2014; Huang et al., 2012; Zikmund-Fisher et al., 2014). Common risk factors for low health literacy, low health numeracy, and low general quantitative literacy were identified in the 2003 NAAL, the 2007 HINTS, and the PIACC studies. These risk factors included non-white minorities, the elderly, those with lower education levels, and those of lower socioeconomic status (Goodman et al., 2013; Huang et al., 2012; Kutner et al., 2006; Smith et al., 2010).

The negative health outcomes associated with low health literacy have been well established throughout the literature (Berkman et al., 2011; DeWalt, Berkman, Sheridan, Lohr, & Pignone, 2004; Omachi et al., 2013). Individuals with low health literacy are more likely to be hospitalized and use the emergency department. They are less likely to use preventative health services and have poorer general health status. The prevalence of chronic diseases and measures of morbidity are higher in individuals with low health literacy. Extensive lists of negative outcomes based on review of the literature are available including such markers as depression, tobacco use, and preteen use of alcohol (Dewalt et al., 2004). Most of the early research focused on overall health literacy but did not specifically consider numeracy as an independent factor (Berkman et al., 2011).

Substantial research has been conducted on the association of health literacy and numeracy to poor outcomes with diabetes such as glycemic control (DeWalt et al., 2004; Garrison et al., 2012; Osborn, Cavanaugh, Wallston, & Rothman, 2010). Zikmund-Fisher et al. (2014) studied adults age 40-70 years old, about half having diabetes, and found that limited
health literacy and numeracy were associated with less sensitivity to hemoglobin A1c levels. Participants were shown laboratory test results in a standardized tabular format and were asked to determine if the results were outside the normal reference range. Of those with higher health literacy and numeracy, 77% could correctly identify the hemoglobin A1c levels as out of range while only 38% of the lower health literacy and numeracy participants could. In the same study, higher health literacy and numeracy were associated with proper determination of when to call the doctor about glycemic control. In a study of Type 1 and Type 2 diabetics, all high school graduates, 33% could not determine total caloric content of a serving from a food label (Garrison et al., 2012). Participants performed even worse calculating the carbohydrate grams of a serving from a nutrition label, with 71% of the Type 2 diabetics and 41% of the Type 1 diabetics unable to do so. The health literacy and numeracy of participants were negatively associated with glycemic control. Hemoglobin A1c levels decreased by 0.35 for Type 1 and 0.52 for Type 2 for every 14% increase in diabetes-related numeracy test score. Osborne et al. (2010) explored self-efficacy as a pathway link between health literacy and numeracy to glycemic control. Health literacy and numeracy were each positively associated with diabetes self-efficacy while diabetes self-efficacy was positively associated with better glycemic control. The results of the study suggested diabetes-related numeracy skills are more strongly associated with glycemic control than general health literacy. Numeracy appeared to be the mediator between health literacy and self-efficacy.

Garcia-Retamero et al. (2015) examined the relationship of subjective (self-reported) and objective (actual) numeracy to subjective (perceptions of physical and mental health) and objective (prevalence of comorbid conditions and prescription medicines) health outcomes. Subjective numeracy was associated with perceptions of health while objective numeracy was
associated with prevalence of comorbid conditions and prescription medicines. Participants with low subjective numeracy had more negative perceptions of health and noted feeling depressed, sad, and anxious more than those with high subjective numeracy (Garcia-Retamero et al., 2015, p. 507). Participants with low objective numeracy had a higher prevalence of comorbidities such as diabetes, chronic obstructive pulmonary disease, or peptic ulcer disease. This study elucidates important relationships between numeracy and health outcomes, but more research is needed focused specifically on numeracy and health outcomes, independent of health literacy.

Numeracy skills are not only needed to manage one’s own health but are also needed by caregivers to properly manage the health of their children (Kumar et al., 2010). In a study of infant (<13 months) caregivers, 99% of the participants had adequate literacy skills but only 17% had greater than 9th grade numeracy skills. Twenty-seven percent of participants were unable to properly dose a liquid prescription medicine using a syringe. Forty-seven percent were unable to determine the correct dose of an over-the-counter, liquid medicine using a dosage chart. Thirty-one percent were unable to determine if they should call their doctor about a fever based on reading a digital thermometer and given an upper range for fever to call about. Fifty-three percent were unable to explain how to make a bottle using concentrated formula that had directions to “mix equal amounts of formula and water” (Kumar et al., 2010, p. 312). Those examples are all basic numeracy skills and are not as complex as interpreting probabilities or understanding risk (Golbeck et al., 2005), yet many caregivers struggled with the basic level tasks.

Aside from following dosing regimens and diabetes management, the most noted influence of numeracy in the literature is on understanding risk and the ability to make informed health decisions (Ahmed et al., 2012; Brust-Renck et al, 2013; French, 2014; Garcia-Retamero &
Hoffrage, 2013; McCaffery et al. 2012; Reyna et al., 2009). Low numeracy can result in distorted perceptions of risk and biases in reasoning and making decisions. The connection of numeracy and decision making is thought to be impacted by metacognitive processes that influence judgment rather than only a function of mathematical computations (Ghazal, Cokely, & Retamero, 2014). The numeracy and decision making correlation goes beyond education-based skills such as basic arithmetic, computational, analytical, and statistical calculations. The relationship is also influenced by emergent decision-based numeracy skills not taught in school such as information seeking, attention, memory, information sensitivity, and affective meaning (French, 2014, p. 99).

Genetic counselors often communicate probabilities and risks to patients (Mann, Mui, Boomsma, & Hasegawa, 2015; Portnoy, Roter, & Erby, 2010). This group of health professionals is trained to translate complicated health information into a format that persons with low health literacy can understand and act upon (Mann et al., 2015). Most often, genetic counselors practice in cancer, prenatal, and pediatric clinical settings; however, less traditional positions, such as in public health, give these health professionals the opportunity to communicate with many different groups. In a study on the impact of numeracy on knowledge in genetic counseling for breast cancer risk, numeracy independently predicted learning of verbally communicated information from videotaped counseling sessions (Portnoy et al., 2010). Genetic literacy was also an independent predictor; however, numeracy and literacy were only moderately associated. Even participants with adequate literacy skills had varied levels of numeracy skills. The results of the study suggested individuals’ levels of numeracy may impact their ability to make informed decisions related to screening and treatment for breast cancer.

**Best practices for communicating risk.** RTI International- University of North
Carolina Evidence-based Practice Center commissioned by AHRQ updated the 2004 systematic review of health literacy interventions and outcomes (Berkman et al., 2011). Interventions for numeracy that improved understanding for individuals with low health literacy included:

- presenting essential information by itself
- presenting essential information first
- presenting health quality information such that the higher number (rather than the lower number) indicates better quality
- using the same denominators to present baseline risk and treatment benefit
- adding icon arrays to numerical presentations of treatment benefit
- adding video to verbal narratives (Berkman et al., 2011, p. ES-7).

These best practices appear throughout the literature, and some studies focus specifically on how to best communicate risk (Fagerlin et al., 2011; Fischhoff et al., 2011; Trevena et al., 2013; Zipkin et al., 2014). The U.S. Food and Drug Administration (FDA) published an evidence-based user’s guide for communicating risks and benefits with three potential goals: share information, change beliefs, and change behavior (Fischhoff et al., 2011). The guide suggested ten ways to present information to promote understanding and better choices. Some of the recommendations echoed those of the AHRQ report. The general practice advice for best practices included:

- provide numeric likelihoods of risks and benefits
- provide absolute risks, not just relative risks
- keep denominators constant for comparisons
- keep time frames constant
- use pictographs and other visual aids when possible
• make the differences between baseline and treatment risks and benefits clear
• reduce the amount of information shown as much as possible
• provide both positive and negative frames
• take care using interpretive labels or symbols to convey the meaning of important information
• test communications prior to use (Fischhoff et al., 2011, pp. 59-61).

Other literature coincides with several of the AHRQ and FDA suggestions but also suggests using plain language in written and spoken communication, using frequencies instead of percentages, and using summary tables when numerous risks and benefits need communicated (Fagerlin et al., 2011). Presenting both frequencies and percentages does not appear to increase comprehension (Trevena et al., 2013). When using frequencies, the best practice of keeping denominators constant must be followed to avoid denominator neglect. For example, in a study people preferred 10-in-13 odds over 9-in-11 odds even though the latter are more favorable, $10/13 \approx 77\%, \; 9/11 \approx 82\%$ (Reyna & Brainerd, 2008). Wilhelms and Reyna (2013, p. 34) added the importance of gist representations of information, “vague, qualitative representations that capture the meaning.” It is suggested that categorical gist drives decisions more than exact numerical values alone. Most experts in the field of risk communication suggest using both numeric and evaluative labels to promote understanding (Trevena et al., 2013).

The FDA referred to the word “pictographs” while AHRQ referred to the words “icon arrays”, but both allude to the same concept of using visuals to depict numeric risk (Berkman et al., 2011; Fischhoff et al., 2011). These words are often used interchangeably in the literature. The concept “less is more” may benefit patient understanding of risk with icon arrays (Zikmund-Fisher, Fagerlin, & Ubel, 2010). When participants in a study were given the choice of icon
arrays with all possible outcomes for survival and mortality or only survival outcomes, participants preferred the latter. In addition to the less complex graphic being the preferred presentation of risk, participants demonstrated a better comprehension of risk with the simpler icon arrays. Gaissmaier et al. (2012) examined risk comprehension differences between participants with low and high graph literacy. The results of the study suggested iconicity of graphics (abstractness versus concreteness) did not make a difference in understanding of risk. Although both low and high graph literacy groups selected graphical representations as more visually attractive than numerical representations, only the high graph literacy group demonstrated better comprehension and recall with the graphical representations. The low graph literacy group demonstrated better comprehension and recall with the numerical representations. Preference of presentation format does not necessarily correlate with better comprehension of information and should be taken into account when communicating risk information to patients.

Although the research suggests using visual aids such as graphical representations, different types of graphs are better for certain numeric information (Brust-Renck et al., 2013; Garcia-Retamero & Cokely, 2013; McCaffery et al. 2012; Wilhelms & Reyna, 2013). Simple bar graphs are recommended to show relative difference between two magnitudes. Pie charts and icon arrays can also demonstrate relative magnitude such as adverse events. Stacked bar graphs prevent denominator neglect and can best show absolute risk. Icon arrays also make the denominator clear and allow representation of relative risk and magnitude. McCaffery et al. (2012) examined whether pictographs or bar graphs were best to communicate small probabilities with denominator of 1000. For smaller numerators, less than 100, pictographs were best understood by the low literacy participants. For larger numerators, more than 100, bar graphs were the best format for low literacy participants. To convey change over time, line
graphs are suggested. Wilhelms and Reyna (2013) explained that line graphs best communicate the gist of the trend. Tables (2x2), Venn diagrams, and Euler diagrams clearly display classes and complex probability.

Furthermore, guidelines have been developed for graphical presentation of quantitative data to ensure the best graph is selected to represent the type of information being depicted (Woller-Carter, Okan, Cokely, & Garcia-Retamero, 2012). Still yet, the media and advertisers often do not follow the guidelines and distort information in graphs to skew judgment and decision making in their favor. These misleading graphical representations can influence even those with adequate numeracy and graph literacy skills, but those with lower numeracy and graph literacy skills are most likely to misinterpret the data.

**Health professional knowledge and use of health literacy and numeracy best practices.** A thorough review of the literature revealed no studies that showed routine, proficient use of health literacy best practices by medical professionals. No studies focused specifically on implementation of numeracy best practices with the exception of research examining how medical professionals communicate risk (Anderson et al., 2011; Han et al., 2014; Neuner-Jehle et al., 2011). Neuner-Jehle et al. (2011) explored how physicians communicate cardiovascular risk in a primary care setting and found verbal qualifiers were used more than other formats. However, the numeracy best practice is to use numerical and visual formats or to use verbal qualifiers with the other formats rather than verbal qualifiers alone (Fagerlin et al., 2011; Fischhoff et al., 2011; McCaffery et al., 2012).

Research on physician assistants and medical school residents has revealed a lack of confidence in communicating with patients with low health literacy (Ali et al., 2014; Green et al., 2014; Karsenty et al., 2013). In an emergency department setting, medical residents did not
demonstrate effective communication tailored to the health literacy of patients (Karsenty et al., 2013). In over half of the encounters, medical acronyms were used. During history taking, residents used medical jargon 66.2% of the time and only provided some further explanation 27.6% of the time. Teach-back, a way of checking for understanding by having patients repeat back information or instructions in their own words (Brega et al., 2015), was not evident in any observed encounter. In additional research on communication techniques used in the emergency department, physicians stated high perceived effectiveness of strategies such as teach-back (64.6%) but reported low routine use (28.4%) (McCarthy, Cameron, Courtney, & Vozenilek, 2012).

Even physicians reporting frequent use of health literacy best practices such as using plain language and teach-back were found to overestimate their actual use of clear communication (Howard et al., 2013). The false perception of clear communication coupled with overestimation of patient literacy level (Bass III, Wilson, Griffith, & Barnett, 2002) could pose barriers to effective communication resulting in lack of patient understanding. Furthermore, if medical professionals often misjudge their ability to communicate, findings from other studies based only on medical professional self-reporting must be interpreted with caution (Cafiero, 2013; Schwartzberg, Cowett, VanGeest, & Wolf, 2007).

Schwartzberg et al. (2007) surveyed physicians, nurses, and pharmacists regarding methods of communication for low health literacy patients and found that 94.7% of the participants reported using plain language. Observational data of those participants in practice to confirm or nullify such a high percentage of plain language use would have added meaning to the research. Similarly, nurse practitioners reported strong intention to use health literacy strategies when communicating with patients (Cafiero, 2013). A validated tool, the Health
Literacy Strategies Behavioral Intention Questionnaire, was used to measure intent. It would be of value to follow the participants and examine if their intent was aligned to their practice. The nurse practitioner intent did not coincide with other research on nurse communication in primary care settings with Type 2 diabetics (Al Sayah, Williams, Pederson, Majumdar, & Johnson, 2014). When patient encounters were observed, nurses used excessive medical jargon and did not routinely use communication loop components. For example, nurses did not check for understanding 81% of the time. Nurses did tend to use less jargon with low health literacy patients, but their use of appropriate communication loops did not differ based on patient health literacy level.

**Health professions curriculum for health literacy and numeracy.** Goal 2 of the National Action Plan to Improve Health Literacy called for educators and licensing and credentialing organizations to “lead the way in changing skills and competencies of professionals” (U.S. Department of Health and Human Services, 2010b, p. 26). Strategies for educators and licensing and credentialing organizations specifically listed incorporating health literacy coursework into curricula, providing training opportunities for students and residents, assessing health literacy competencies for licensing, and developing continuing education requirements in health literacy for all health professions (U.S. Department of Health and Human Services, 2010b, p. 28). The IOM report on health literacy provided similar recommendations: “Professional schools and professional continuing education programs in health and related fields, including medicine, dentistry, pharmacy, social work, anthropology, nursing, public health, and journalism, should incorporate health literacy into their curricula and areas of competence” (Nielsen-Bohlman et al., 2004, p. 161). The Affordable Care Act of 2010 also included a section requiring health literacy in health professional training (Koh et al., 2012). Yet
health literacy competencies have not been fully integrated into all health professions curricula (Ali, 2012; Coleman, 2011; Coleman & Appy, 2012; Toronto & Weatherford, 2015).

The Liaison Committee on Medical Education (LCME) of the American Association of Medical Colleges (AAMC) requires medical school curriculum to include health literacy training (Ross, Lukela, Agbakwuru, & Lypson, 2013). Coleman and Appy (2012) studied a sample of U.S. medical schools in 2010 and found 72.1% reported including health literacy in the required curriculum, but the average time dedicated to health literacy teaching was only three hours. This lack of integration of health literacy into the curricula is also evident in nursing and pharmacy programs (Cailor & Chen, 2015; McCleary-Jones, 2012; Schwartzberg et al., 2007; Toronto & Weatherford, 2015).

Ali (2012) surveyed faculty of internal medicine residency programs and found that less than half of the programs included health literacy as part of their formal curriculum. Often the term “health literacy” is not specified in competencies for communication (Henry et al., 2013). One article suggested 12 evidence-based competencies for communication in residency programs but never uses the term “health literacy.” It generically stated “Ability to communicate treatment plans” (Henry et al., 2013, p. 397) but explicitly listed teach-back as an example of evidence of that competency. Most terminology for competencies from The Accreditation Council for Graduate Medical Education (ACGME) only refer to communication and interpersonal skills, but the Family Medicine Milestone Project, a joint initiative between ACGME and The American Board of Family Medicine, specifically lists “health literacy” one time in the document outlining the competencies for family medicine physicians (Family Medicine Milestone Project, 2014, p. 85). The document includes describing risk as a competency but does not elucidate how or what best practices to use. Englander et al. (2013) suggest 58 competencies in 8 domains for
physicians based in part on the ACGME/American Board of Medical Specialties competencies in 6 domains. The competencies include empowering patients for shared decision making as well as communicating “effectively with patients, families, and the public, as appropriate, across a broad range of socioeconomic and cultural backgrounds” (Englander et al., 2013, p. 1091). However, the terms “communicating risk,” “health literacy,” and “numeracy” are not used.

A plethora of recommendations are available for health literacy training interventions for health professions including pharmacy, nursing, and medicine (Bloom-Feshbach et al., 2015; Cailor & Chen, 2015; Coleman & Fromer, 2015; Coleman et al., 2013; Green et al., 2014; Ha & Lopez, 2014; Pagels et al., 2015; Roberts et al., 2012). The teaching and instructional methods vary as does the recommended time frame for the learning experiences within the curricula. Coleman and Fromer (2015) evaluated the impact of a 3.5-hour health literacy training intervention for physicians and other health professionals at different points in their education. The intervention included didactic instruction followed by an experiential workshop and resulted in increased “knowledge, perceived skill, and intended behavior” (Coleman & Fromer, 2015, p. 388) and varied by profession and by years of experience for physicians.

Green et al. (2014) developed a health literacy training program for internal medicine residents including didactic instruction, standardized patient encounters, and feedback from videotaped sessions. An evaluation of the curriculum showed it to be well-received by learners and to improve “knowledge, attitudes, and skills regarding health literacy” (Green et al., 2014, p. 76). Bloom-Feshbach et al. (2015) evaluated a health literacy workshop and OSCE (Objective Structured Clinical Examinations) for fourth year medical students and reported the program may lead to improved health outcomes and patient safety. Pagels et al. (2015) developed a health
literacy curriculum for family medicine residents that included didactic lecture and OSCE to address health literacy knowledge, communication skills, and using an interpreter. Promotoras, community health workers, served as standardized patients in the OSCEs. The authors reported an increase in health literacy knowledge, acceptable OSCE scores, high use of teach-back (77.8%), and using an interpreter effectively (77.8%). However, they also discussed that a one-time training is not enough to address the problem of low health literacy and suggested customized training based on specific patient populations. A 6-week health literacy curriculum was designed and evaluated for a third year family medicine clerkship (Roberts et al., 2012). Through didactic instruction, simulation with standardized patients, and an online discussion board, student knowledge and application of health literacy best practices improved. Risk communication, however, was not included as an outcome measure.

Lopez (2014) recommended case-based learning for year three pharmacy students in addition to the current one-hour didactic session and three-hour lab during the first year of training. Cailor and Chen (2015) suggested integrating health literacy concepts into multiple courses in the first year of a pharmacy curriculum through didactic instruction, independent readings, and active-learning experiences. Ross et al. (2013) explored the impact of a single health literacy activity embedded in the longitudinal curriculum for sociocultural topics. The second year medical students identified low health literacy as a barrier to care but did not recognize it as a social determinant of health, evidence that more explicit focus on health literacy is needed in the curriculum. Coleman, Peterson-Perry, and Bumsted (2016) further demonstrated that a one-time, didactic format training in health literacy is not enough to improve perceived knowledge and intent to use best practices. Although post-test scores were significantly higher than pre-test scores immediately following the health literacy training of first year medical
students, a drastic drop was seen in perceived knowledge and planned behaviors after 12 months. The improvement between pre- and post-tests of second year medical students receiving health literacy training was less than that for first year students. The authors concluded a “longitudinal or integrated format” (p. 53) should be used for health literacy training within the medical school curriculum.

General health literacy best practices for communication such as plain language and using teach-back are included in the research for interventions and curriculum, but numeracy best practices and ways to communicate risk are not. For example, the communication skills of medical students improved with a training intervention of translating medical documents into plain language (Bittner, Jonietz, Bittner, Beickert, & Harendza, 2015). A study by Bittner et al. (2015) focused on eliminating medical jargon but did not mention using numeracy best practices. In a consensus study suggesting curriculum changes to adapt health literacy competencies, only 1 of 32 practice competencies is related specifically to health numeracy and is not delineated beyond “The individual routinely conveys numeric information, such as risk, using low numeracy approaches, such as through examples, in oral and written communication” (Coleman et al., 2013, p. 95).

Shared Decision Making

Over four decades ago, the concept “sharing of decision making” was first used by Veatch (1972) referring to ethical practices for patient-provider communication in medicine. However, it was another quarter of a century until “shared decision making” appeared again in research literature (Stiggelbout, Pieterse, & De Haes, 2015). Now it is called for in federal initiatives (U.S. Department of Health and Human Services, 2010a) but is not implemented routinely in practice for various possible reasons (Garcia-Retamero, Wicki, Cokely, & Hanson,
Stiggelbout et al. (2015) explained shared decision making as part of two fields, both in medical ethics and in health services research. After examining different definitions and frameworks for shared decision making, Stiggelbout et al. (2015) offered four specific steps for the process:

1. The professional informs the patient that a decision is to be made and that the patient’s opinion is important;
2. The professional explains the options and the pros and cons of each relevant option;
3. The professional and patient discuss the patient’s preferences; the professional supports the patient in deliberation;
4. The professional and patient discuss patient’s decisional role preference, make or defer the decision, and discuss possible follow-up (p. 1173).

Step two could involve expressing numerical risk and benefit to a patient. Health professionals would need to utilize effective best practices for risk communication. Even if the four steps are followed, some patients will still prefer for the professional to make the choice for them. Both patients and providers need experience and training for effective partnerships and shared decision making.

Elwyn et al. (2012) suggested a three-step model for shared decision making in clinical practice: “a) introducing choice, b) describing options…, c) helping patients explore preferences and make decisions” (p. 1361). The authors explain that describing options includes discussion of the “harms and benefits” and requires use of “effective risk communication” (p. 1364). Furthermore, Elwyn et al. (2012) pointed out the challenge posed by uncertainty in medicine and the need to share any unpredictable outcomes with patients. Han (2013) described three sources
of uncertainty in clinical evidence: probability, ambiguity, and complexity. Communicating this uncertainty in shared decision making leads to conceptual, methodological, and ethical problems. In his research, Han consistently focused on how to communicate risk to patients but overlooks the possibility of low statistical literacy of physicians as a barrier (Han, 2013; Han et al., 2014).

The bulk of literature suggests health literacy and numeracy of patients along with physician numeracy impact shared decision making (Garcia-Retamero et al., 2014; Goggins et al., 2014; Hanoch, Miron-Shatz, Rolison, Omer, & Ozanne, 2014; Légaré & Witteman, 2013; Smith, Nutbeam, & McCaffery, 2013; Yin et al., 2012). In an inpatient setting of individuals with cardiovascular disease, the preferences for involvement in decision making varied based on patient characteristics (Goggins et al., 2014). Patients with lower health literacy and numeracy levels preferred less involvement in decision making than those with higher health literacy and numeracy. Similar results were found with patients at risk for cancer (Hanoch et al., 2014) with high numeracy positively related to patient desire to have an active role in decision making. Low health literacy proved to be a barrier to shared decision making for parents of children in a pediatric, primary care setting (Yin et al., 2012).

The complex demands on the patient for shared decision making reflect the three aspects of health literacy as described by Nutbeam: functional, interactive, and critical health literacy (Smith et al., 2013). The three aspects are defined as

- Functional health literacy is the ability to obtain, understand and use factual information on health risks and on how to use the health system.

- Interactive health literacy is defined as the ability to extract health information and derive meaning from different forms of communication, and to apply new information to changing circumstances.
Critical health literacy reflects the literacy and numeracy skills that support critical reflection on information or advice received, including recognition of the influence of wider social determinants of health (Smith et al., 2013, p. 1013).

Nutbeam’s model has been applied to multiple situations including addressing issues of social determinants of health. Health literacy often is related to other social determinants of health and is negatively correlated. Patients with low health literacy or numeracy are less likely to be successful with these complex demands required for shared decision making. For example, Smith et al. (2013) found that low literacy participants did not understand graphical risk diagrams in decision aids specifically designed to make explaining risk clearer. Low literacy participants tended to have less self-confidence in their health decision making capabilities, consistent with other research.

Patient preference for involvement is only one side of the equation for shared decision making. Physician preference plays an equally important role and could possibly be the driving force behind whether shared decision making is attempted (Garcia-Retamero et al., 2014). Garcia-Retamero et al. (2014) studied an international sample of surgeons from 60 countries to determine factors associated with preferences and actual implementation of shared decision making. With the Berlin Numeracy Test as the measure for numeracy level, low numeracy physicians were less likely to participate in shared decision making even though they expressed value in the collaborative process with patients. This provided evidence that both patients’ and health professionals’ willingness to participate in shared decision making is influenced by numeracy. The authors suggest communicating risk using numerical information is a barrier to effective patient-provider collaboration.

Although studies have revealed that informed consent, a part of shared decision making,
requires communicating risk to patients, little research has been conducted on how physicians convey statistical information (Gaissmaier, Anderson, & Schulkin, 2014). A study of obstetricians and gynecologists found that most did not give full information nor were transparent with patients. Only one quarter of physicians in the study provided both complete and transparent information, and most did not routinely follow numeracy best practices for communicating risk. The authors suggested a lack of statistical knowledge could have been an underlying cause of physicians providing misinformation.

**Statistical Numeracy**

**Health professional statistical numeracy.** The focus of health numeracy is generally on patients rather than health professionals. Research has examined the biostatistical knowledge of health professionals and consistently found deficits (Anderson et al., 2014; Berwick et al., 1981; Susarla & Redett, 2014; Weiss & Samet, 1980; Windish et al., 2007; Wulff et al., 1987). Although biostatistical knowledge is needed to interpret the results of clinical studies (Anderson et al., 2014; Narayanan, Nugent, & Nugent, 2015; Johnson et al., 2014), most research overlooks the need for basic statistical literacy for judgment and decision making (Ghazal et al., 2014). Wegwarth and Gigerenzer (2011) explained that most doctors struggle with medical statistics even in their own specialty, resulting in lack of understanding of survival rates, relative risk reductions, and the benefits and risks of screening. Furthermore, they stated statistical illiteracy in doctors prevents effective communication of risk needed for shared decision making. Caverly, Prochazka, Binswanger, Kutner, & Matlock (2012) suggested physician numeracy could affect both health decision making as well as risk communication to patients. Certainly, low statistical literacy rather than biostatistical knowledge could be to blame for errors in physician judgment and decision making (Ghazal et al., 2014; Trickey et al., 2014).
A review of 49 articles from leading medical journals such as the *New England Journal of Medicine*, the *Journal of the American Medical Association*, and the *Annals of Internal Medicine* revealed the need for physicians to understand complex statistical concepts (Narayanan et al., 2015). Statistical summary measures and models included basic descriptive statistics (i.e. percentages, mean, median, range, and standard deviation), group comparisons (i.e. chi-squared, t test, and ANOVA), regression analysis, ratios (i.e. confidence intervals, odds, and hazards ratio), and test analysis (i.e. positive/negative predictive value). Narayanan et al. (2015) determined the level of statistical mastery of terms used in the articles and reported 13.4% as introductory level, 22.3% as advanced undergraduate level, 34.8% as master’s level, and 29.5% as doctorate level. Yet these journals are routinely used in undergraduate medical curriculum to teach evidence-based medicine to students that have not had introductory statistics courses.

In an international sample of surgeons from 60 countries that were administered the Berlin Numeracy Test, 50% of participants scored in the lowest of four levels of statistical literacy (Garcia-Retamero et al., 2014). The test measures statistical numeracy and risk literacy, not specifically biostatistical knowledge. Another study used the Schwartz Numeracy Scale as the instrument to measure objective numeracy in a sample of obstetrician-gynecologists (Anderson et al., 2011). The three-question assessment did not include biostatistics or even what most would consider difficult statistics but rather included converting a frequency to a percent, converting a percent to a frequency, and estimating how many heads would be in 1000 flips of a coin. However, only 66.1% of participants answered all three questions correctly. A different sample of obstetrician-gynecologists was assessed using a combination of measures: the Schwartz Numeracy Scale, the Lipkus Numeracy Scale, and the Obstetrician-Gynecologist Statistical Literacy Questionnaire (OGSLQ) (Anderson et al., 2014). Although the sample of
physicians performed somewhat better on the Schwartz Numeracy Scale than in the aforementioned study with 78% answering all three questions correctly, overall performance on the measures of numeracy indicated physicians were lacking in statistical literacy needed for making treatment decisions.

Studies have demonstrated physicians misunderstanding test results (Wegwarth et al., 2012) and misinterpreting risk data (Caverly et al., 2015). Physicians often struggle with positive predictive value, being able to estimate the probability of a disease based on a positive test result (Gigerenzer, 2014). Gigerenzer (2014) found only 1 out of 24 physicians accurately estimated the probability of a patient having colorectal cancer given a positive fecal occult blood screen test. A 1978 study on positive predictive value was replicated in 2013 with a sample of attending physicians, house officers, medical students, and a retired physician (Manrai, Bhatia, Strymish, Kohane, & Jain, 2014). The group was asked the same question used in 1978 that previously showed an overestimation of positive predictive value,

If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming you know nothing about the person’s symptoms or signs? (Manrai et al., p. 991)

Similar to the original study, less than a quarter of respondents answered the question correctly. The correct answer was 2%, but the most common answer given was 95%, a vast overestimate.

Anderson et al. (2011) reported physician numeracy level impacted how Down syndrome screening test results were communicated to patients. Physician numeracy influences attitudes toward cancer screening with higher numeracy inversely related to favoring screening (Caverly et al., 2012). The attitude toward cancer screening could impact how physicians communicate
risks and benefits to patients. Additionally, physicians, like the general population, have problems with relative risk information despite it being the most common method of reporting efficacy in clinical trials both in scientific papers and the media (Marcatto, Rolison, & Ferrante, 2013). When physicians were presented with a hypothetical clinical trial for an antihypertensive drug, their interpretations of the drug’s efficacy were skewed by presentation format, relative or absolute risk reduction. Physicians perceived relative risk reduction as more effective than absolute risk even when presented with baseline risk information. Marcatto et al. (2013, p. 29) recommended reporting clinical outcomes in terms of absolute risk reductions due to the biasing effects of relative risk reduction formats. Some of the best practices for communicating risk to patients may also benefit physician understanding of statistical information such as using visual representations and using natural frequencies instead of probabilities (Gargia-Retamero & Hoffrage, 2013).

**Health professions student statistical numeracy.** Clinicians are required to understand and interpret statistical information to practice evidence-based medicine (Arnold, Braganza, Salih, & Colditz, 2013; Baghi & Kornides, 2013; Friederichs, Ligges, & Weissenstein, 2014; Narayanan et al., 2015; Trickey et al., 2014), yet the attitudes and knowledge of statistics of health professions students is often not favorable (Beurze et al., 2013; Bookstaver et al., 2012; Hannigan et al., 2014; Johnson et al., 2014; Kiekkas et al., 2015; Zhang, 2012). Postgraduate year one pharmacy residents demonstrated poor biostatistics knowledge in an online survey, consistent with previous results (Bookstaver et al., 2012). Both attitudes and confidence ratings had statistically significant, positive correlations to knowledge scores. Kiekkas et al. (2015) examined undergraduate nursing students’ attitudes toward statistics, pre and post a biostatistics course, using The Survey of Attitudes Toward Statistics (SATS)-36 scale. Consistent with
previous research, nursing students expressed anxiety toward statistics, but attitudes improved as a result of the coursework. Additionally, overall SATS-36 scores were correlated with examination scores, but the statistically significant, positive relationship was weak.

The attitudes and statistical knowledge of healthcare major students in a graduate level, introductory statistics course were assessed pre- and post-instruction (Baghi & Kornides, 2013). Over a third of the student sample were registered nurse practitioners or clinical nurse specialists. An instrument was developed to measure six domains of attitudes toward statistics with four focused on specific self-reported knowledge components, one on utility, and one on self-confidence. Actual statistical knowledge was measured by a five-subscale proficiency assessment. Contrary to most of the literature, students in the study expressed largely positive attitudes toward statistics on the pre-test. Both attitudes and knowledge scores significantly improved from pre- to post-test except for the domain of utility. The initial utility scores indicated students recognized the value of statistics for their profession leaving little room for improvement. Level of proficiency and attitudes toward statistics were associated, and improved understanding may have impacted student attitudes. Baghi and Kornides (2013) ascertained initial assessment of student attitudes toward statistics should be used to guide training and interventions. Targeting students with negative attitudes and providing them additional training could result in greater use of evidence-based practice upon graduation.

Thompson, Wylie, Mulhern, and Hanna (2015) developed path analysis models based on affective, demographic, and educational predictive variables effects on numeracy performance of psychology, nursing, and medicine undergraduate students. They discussed previous research associating math anxiety and math performance as well as the relationship between math anxiety and the variables of gender and age. The Fennema-Sherman Mathematics Attitude Scale
measured mathematics attitude and included four subscales: math anxiety, motivation, usefulness, and confidence (Thompson et al., 2015, p. 134). Math ability was measured by a 20 question, 32 problem numerical knowledge test that included six categories: decimal and fraction calculations, algebraic reasoning, graphs, ratios and proportions, probability and sampling, and estimation (Thompson et al., 2015, p. 135). Although the numeracy test did not specifically measure statistical knowledge alone, the authors claimed the instrument was shown to significantly predict psychology statistics exam scores. The researchers found differences among the disciplines with psychology and nursing student performance most predicted by math anxiety while medical student performance was most predicted by motivation. Path analysis revealed that for all disciplines studied, confidence only had an indirect effect on math performance.

Beurze et al. (2013) explored medical student math anxiety, specifically, statistics anxiety, as a possible barrier to student success in research methodology coursework. The Statistical Anxiety Rating Scale (STARS) scores of first and second year medical students had only a small effect on course performance, and only moderate statistics anxiety was reported. Although a negative correlation was found between math performance in high school and statistics anxiety, there was no relationship between previous statistics education and statistics anxiety in the study sample.

Hannigan et al. (2014) and Zhang et al. (2012) used the SATS to investigate the attitudes toward statistics of entry-level graduate medical students and postgraduate medical students, respectively. Past math performance was a predictor of attitudes for both cohorts with a significant positive relationship. Although the majority, 85%, of entry-level graduate medical students had previously taken a quantitative course, only 24% would voluntarily take a statistics
course (Hannigan et al., 2014). The postgraduate medical students completed the statistics attitudes assessment before and after a statistics course (Zhang et al., 2012). There was a positive relationship between course achievement and statistics attitudes, and overall, attitudes toward statistics declined after taking the course.

The importance of statistical numeracy of health professionals includes being able to interpret evidence-based medical information and understand risk (Johnson et al., 2014; Narayanan et al., 2015). Numeracy of medical students and surgical residents, as measured by the Schwartz-Woloshin three-item numeracy tool, was an independent predictor of risk comprehension. Only 69% of the participants answered all three questions correctly, and those with inadequate numeracy had a seven-fold increased change of misunderstanding risk than their numerate counterparts. Neither confidence levels nor training level were statistically correlated to risk comprehension; nor was there a statistically significant relationship between confidence level and training level. Johnson et al. (2014, p. 211) concluded the innumeracy of medical students leads to “misunderstanding and miscommunication of risk” that “could potentially affect patient safety and care.”

In addition to lack of knowledge and negative attitudes toward statistics, often medical students do not see the relevancy of statistics to their profession until they actually begin practicing (Freeman, Collier, Staniforth, & Smith, 2008; Miles, Price, Swift, Shepstone, & Leinster, 2010). Although only 40% of practicing physicians (n=130) reported finding probability and statistics relevant to their future when they were in undergraduate medical training, 73% of the cohort reported that these were, in fact, needed in clinical practice (Miles et al., 2010). Many of the group surveyed explained the teaching methods did not stress the
relevancy to their future role and that statistics should be taught with reference to clinical practice and research rather than simply in an abstract mathematical context.

**Medical school curriculum for statistics.** Thomas et al. (2015, p. 1) explained medical education curriculum must adapt as the needs “of patients, medical practitioners, and society change” and provide a list of contemporary demands of medical education. The authors included the competency of practicing evidence-based medicine and using scientific evidence to make appropriate clinical decisions. Statistical literacy is required to effectively practice evidence-based medicine (Johnson et al., 2014; Trickey et al., 2014; Wegwarth & Gigerenzer, 2011). An exploration of a leading medical journal, the *Journal of the American Medical Association*, over a 20-year period revealed an increase in the use and complexity of statistical information (Arnold et al., 2013). In addition to the traditionally reported descriptive statistics and epidemiological data, more advanced statistical methods were found in the later decade of the timeframe studied. Randomized controlled trials, meta-analyses, sensitivity analyses, survival analyses, and multiple regression increased in frequency over the 20-year period. Arnold et al. (2013) concluded medical school curriculum should be revised to include more training in statistics.

The literature established in the 1980’s that physicians lack statistical literacy (Berwick et al., 1981; Weiss & Samet, 1980; Wulff et al., 1987), and more current research has confirmed the problem still exists (Anderson et al., 2014; Susarla & Redett, 2014; Windish et al., 2007). Medical school curricula fail to adequately prepare physicians to understand health statistics and practice evidence-based medicine (Johnson et al., 2014; Wegwarth & Gigerenzer, 2011). The AAMC has not promoted initiatives to address statistical literacy in U.S. medical schools through its accreditation process, and the American Board of Internal Medicine (ABIM) has not placed a significant emphasis on statistical knowledge through its certification testing (Wegwarth
The ABIM certification exam does not include risk communication and only covers medical statistics in 3% of the questions. Some medical schools require biostatistics or epidemiology coursework, but the requirement is not consistently seen across the U.S. Wegwarth and Gigerenzer (2011) did not find risk communication embedded in any of the medical school curricula they reviewed.

The United States Medical Licensing Examination (USMLE) made changes in 2015-2016 to include biostatistics and epidemiology in both Step 2 and Step 3 exams (USMLE, 2015). The USMLE is a three-part exam required for medical licensure in the U.S. Step 2 consists of clinical knowledge and clinical science and is focused on application of medical knowledge, skills, and understanding needed to provide supervised patient care. This step stresses health promotion and disease prevention. Step 3, the final examination for licensing, is focused on application of medical knowledge and understanding of biomedical and clinical science needed for unsupervised patient care. This step stresses patient management in ambulatory settings.

Medical schools often design curriculum based upon competencies found in exams such as the USMLE and the ABIM certification exam (Dexter, Koshland, Waer, & Anderson, 2012; Khan et al., 2015). Some medical schools have reported curriculum changes to reflect more attention on statistical literacy such as biostatistics and understanding risk (Wegwarth & Gigerenzer, 2011). The changes in curriculum have not kept up with the changes in the statistical reporting in the literature or the need for statistical literacy to practice evidence-based medicine (Arnold et al., 2013).

Recent research has explored methods for improving how medical statistics are taught such as practical, applicable, contextual instruction or blended learning over traditional instruction (Evans et al., 2016; Freeman et al., 2008; Masel, Humphrey, Blackburn, & Levine,
Almost a decade ago, a new style of teaching statistics to undergraduate medical students was investigated as a possible solution to the issues of students finding statistics irrelevant and performing poorly (Freeman et al., 2008). Using different teaching modalities such as videos and animations with real world examples in context as the intervention versus traditional instruction, a difference was seen in statistics knowledge and attitudes between groups of students. More students in the intervention group were able to define $p$-value and confidence interval and also expressed agreement regarding statistics being relevant to their medical career.

Asserting a high level of statistical numeracy is needed for informed decision making and that many health-major students lack the ability to reason about chance events in probability, Masel et al. (2015) evaluated an undergraduate probability and statistics course focused on evidence-based medicine. The course was a substitute for a biostatistics class but also met the criteria to be a substitute for a bioethics or a science and society course. Relying on intrinsic motivation of students to learn about evidence-based medicine, the main pillar of the course was the randomized control trial. This allowed for the teaching of important statistical concepts but all in context. Additional interest was gained by using methods like playing cards and rolling dice to understand probability and uncertainty. Pre- and post-tests revealed positive outcomes in student learning of quantitative information as well as an increase in Attitudes Toward Statistics scores. “Flipped” or blended learning environments have been evaluated with positive results supporting the option of alternative environments for teaching medical statistics (Evans et al., 2016; Milic et al., 2016). Still yet, individual learning styles of students were evident as deciding factors for preference of material delivery format. Studies have been conclusive that relevancy to practice needs emphasis regardless of instructional modality utilized.
Beyond the concern of medical school curricula not including enough coursework on biostatistics and understanding risk, basic statistics courses are not taught in medical school (Wegwarth & Gigerenzer, 2011). Medical students have varied background experience in statistics instruction (Beurze et al., 2013; Evans et al., 2016; Freeman et al., 2008). Students may enter medical school without prior coursework in statistics because statistics is often not a pre-matriculation requirement (see Table 1). An examination of the admission requirements for 27 medical schools in the U.S. revealed only three required a course in statistics as a prerequisite. Twelve of the 27 medical schools did not require a course in statistics or even list it as a recommendation. Twelve did not require a course in statistics but did either recommend it or list it as an option alongside other math courses such as calculus. However, Manrai et al. (2014) recommended premedical education coursework focus more on statistics than calculus, claiming statistical reasoning is actually needed to practice medicine while calculus is not. Additionally, they suggested medical training beyond undergraduate include application of medical statistics.

Ivy league medical schools at Yale University (2016) and Harvard University (2016) both stress the importance of a liberal arts education in no particular undergraduate field but differ drastically in their pre-matriculation requirements for math. Yale focuses prerequisite learning in the sciences, specifically requiring biology (or zoology), general chemistry, organic chemistry, biochemistry, and physics but does not require any math coursework. Harvard requires the same sciences with the exception zoology but also requires a full year of math comprised of calculus and statistics. Additionally, the Harvard application requirements recommend biostatistics and require expository writing.

The two medical schools for this study share common science requirements but differ in their math prerequisites. Both schools require biology, general chemistry, organic chemistry,
biochemistry, physics, and genetics for pre-matriculation. University A requires a course in statistics but does not list any other math. University B requires one mathematics course but lists “not including statistics;” however, University B strongly recommends a course in statistics. Additionally, both schools require humanities, social sciences, and English courses.
Table 1

Statistics Course Admission Prerequisite Requirements of 27 U.S. Medical Schools

<table>
<thead>
<tr>
<th>Institution</th>
<th>Required</th>
<th>Not Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boston University School of Medicine (2016)</td>
<td>X***</td>
<td></td>
</tr>
<tr>
<td>Brown University- Alpert (2016)</td>
<td>X***</td>
<td></td>
</tr>
<tr>
<td>Case Western Reserve University (2016)</td>
<td>X*</td>
<td></td>
</tr>
<tr>
<td>Columbia University (2016)</td>
<td>X***</td>
<td></td>
</tr>
<tr>
<td>Dartmouth College (2016)</td>
<td>X**</td>
<td></td>
</tr>
<tr>
<td>Duke University (2016)</td>
<td>X*</td>
<td></td>
</tr>
<tr>
<td>Harvard University (2016)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Indiana University at Indianapolis (2016)</td>
<td>X***</td>
<td></td>
</tr>
<tr>
<td>Johns Hopkins University (2016)</td>
<td>X***</td>
<td></td>
</tr>
<tr>
<td>Mayo Medical School (2016)</td>
<td>X***</td>
<td></td>
</tr>
<tr>
<td>Oregon Health &amp; Science University (2016)</td>
<td>X*</td>
<td></td>
</tr>
<tr>
<td>University of Arkansas for Medical Sciences (2015)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>University of California at San Francisco (2016)</td>
<td>X***</td>
<td></td>
</tr>
<tr>
<td>University of Colorado at Denver (2016)</td>
<td>X***</td>
<td></td>
</tr>
<tr>
<td>University of Kansas Medical Center (2016)</td>
<td>X***</td>
<td></td>
</tr>
<tr>
<td>University of Illinois (2015)</td>
<td>X*</td>
<td></td>
</tr>
<tr>
<td>University of Iowa- Carver (2016)</td>
<td>X**</td>
<td></td>
</tr>
<tr>
<td>University of Massachusetts at Worcester (2016)</td>
<td>X*</td>
<td></td>
</tr>
<tr>
<td>University of Michigan at Ann Arbor (2016)</td>
<td>X*</td>
<td></td>
</tr>
<tr>
<td>University of Minnesota (2016)</td>
<td>X*</td>
<td></td>
</tr>
<tr>
<td>University of North Carolina at Chapel Hill (2016)</td>
<td>X***</td>
<td></td>
</tr>
<tr>
<td>University of Pennsylvania (2016)</td>
<td>X*</td>
<td></td>
</tr>
<tr>
<td>University of Rochester (2016)</td>
<td>X*</td>
<td></td>
</tr>
<tr>
<td>University of Utah (2016)</td>
<td>X***</td>
<td></td>
</tr>
<tr>
<td>University of Washington (2016)</td>
<td>X***</td>
<td></td>
</tr>
<tr>
<td>University of Wisconsin at Madison (2016)</td>
<td>X</td>
<td>X***</td>
</tr>
<tr>
<td>Yale University (2016)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total 3 24

*Recommended in admission coursework requirements

**Listed as an option along with other math coursework

***Not mentioned in admission coursework requirements
Summary

The prevalence of inadequate health literacy and specifically health numeracy in the U.S. has been well established (Goodman et al., 2013; Kutner et al., 2006; Smith et al., 2010) as have the negative outcomes associated with low health literacy and health numeracy (Berkman et al., 2011; Garcia-Retamero et al., 2015; Omachi et al., 2013). Although some curricula have been developed to improve health professional communication with patients with low health literacy, the training is not consistently implemented and does not adequately address communicating risk (Coleman & Fromer, 2015; Coleman et al., 2013; Green et al., 2014; Henry et al., 2013). Furthermore, physicians may lack the needed statistical literacy to effectively communicate risk and benefit information to patients for informed decision making (Ahmed et al., 2012; Anderson & Schulkin, 2011; Caverly et al., 2015; Garcia-Retamero & Hoffrage, 2013; Han et al., 2014; Neuner-Jehle et al., 2011). The literature shows a need to improve communication of risk to patients (Berkman et al., 2011; Han et al., 2014) as well as the problem of biostatistical illiteracy of physicians (Anderson et al., 2014; Susarla & Redett, 2014). However, the underlying issue of a gap in the curriculum for understanding risk and general statistics has not been researched. The root cause of poor communication of risk to patients could be inadequate basic statistical literacy of physicians due to a gap in the curriculum that includes biostatistics but not foundational statistics. More research is needed to explore the relationship between medical students’ risk literacy and attitudes toward statistics to better understand and address gaps in the curriculum for this competency (Anderson et al., 2011; Levy, Ubel, Dillard, Weir, & Fagerlin, 2013).
CHAPTER THREE: METHODS

Overview

This chapter begins with an explanation of the design of the study followed by statement of the research question and corresponding null hypotheses. Details regarding the participants and setting are provided including gender, race, and ethnicity demographics. The instruments used to measure the predictor and criterion variables are described in depth with validity and reliability reported. Procedures of the study are then presented and the data analysis methods identified. Initial data screening and assumption tests required are specified.

Design

This quantitative, non-experimental study used a predictive, correlational design (Warner, 2013) as it investigated the relationship between three predictor variables and one criterion variable. Regression analysis with multiple predictor variables allows assessment of how well the combination of variables can predict the criterion variable. Additionally, multiple regression analysis shows how much each individual predictor variable contributes to the variance while controlling for the other predictor variables. The three predictor variables were score on the Attitude Toward the Field of Statistics subscale, score on the Attitude Toward the Course of Statistics subscale, and successful completion of a college level statistics course. The criterion variable was score on the Berlin Numeracy Test.

The ATS subscales measure “two distinct aspects of student attitudes toward statistics” (Wise, 1985, p. 404). The Field subscale measures the attitude toward the usefulness of statistics either in general or specifically in the student’s field of study. The Course subscale measures the attitude toward a statistics course. Successful completion of a college level statistics course is a self-report measure, yes or no, to completing a college level statistics course with a grade of C or
better. The Berlin Numeracy Test measures statistical numeracy and risk literacy in educated and highly-educated samples (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012). Statistical numeracy is “an understanding of the operations of probabilistic and statistical computation, such as comparing and transforming probabilities and proportions” (Cokely et al., 2012, p. 25). Risk literacy is the “the ability to accurately interpret and act on information about risk” (Cokely et al., 2012, p. 26).

**Research Question**

RQ1: How accurately can risk literacy as measured by the Berlin Numeracy Test be predicted from a linear combination of score on the Attitude Toward the Field of Statistics subscale, score on the Attitude Toward the Course of Statistics subscale, and successful completion of a college level statistics course?

**Hypotheses**

The null hypotheses for this study are:

H₀₁: There is no statistically significant predictive relationship between risk literacy as measured by the Berlin Numeracy Test and the linear combination of score on the Attitude Toward the Field of Statistics subscale, score on the Attitude Toward the Course of Statistics subscale, and successful completion of a college level statistics course.

H₀₂: There is no statistically significant predictive relationship between risk literacy as measured by the Berlin Numeracy Test and score on the Attitude Toward the Field of Statistics subscale.

H₀₃: There is no statistically significant predictive relationship between risk literacy as measured by the Berlin Numeracy Test and score on the Attitude Toward the Course of Statistics subscale.
**H04:** There is no statistically significant predictive relationship between risk literacy as measured by the Berlin Numeracy Test and the successful completion of a college level statistics course.

**Participants and Setting**

The target population was first year medical students in the U.S. with a convenience sample of 327 medical students at two academic health institutions in two states, one in the South Central region and one in the Pacific Northwest. These two medical colleges were selected based on access through connections with colleagues at the institutions that value research pertaining to health literacy and communication with patients. Warner (2013, p. 570) recommends a sample size of $N \geq 50 + 8k$ for tests of multiple $R$ and $N \geq 104 + k$ for tests of significance of individual predictors, where $k$ is the number of predictor variables. To exceed the recommended sample size of $N \geq 107$ for medium effect size, statistical power of .7, and alpha of .05, 150+ responses were collected. University A and University B have 175 and 157 first year medical students, respectively, and the response rates for this study were 98.2% (N=172) and 98.7% (N=155). Participants were 44.0% male, 56.0% female, and 0% other gender; however, the gender breakdown varied noticeably between data collection sites with University A having 48.8% male and 51.2% female and University B having 38.7% male and 61.3% female. Participants self-reported as 7.1% Hispanic/Latino/a, or Spanish origin, 3.1% Black/African American, 0.9% American Indian/Native Alaskan, 0.9% Native Hawaiian/Pacific Islander, 81.7% White, 13.0% Asian, and 4.3% other race. Surveys were administered during classes that enroll all first year medical students at each institution, and participation was voluntary. The Associate Dean at University A and the Chair of the Curriculum Committee at University B were asked by the researcher with support from other faculty to allow and support administering the
surveys during class. University A requires a college level statistics course but no other math as a pre-matriculation requirement. University B requires a college level math class other than statistics but also strongly recommends statistics.

Instrumentation

The Berlin Numeracy Test (see Appendix A for instrument) was used to measure the criterion variable, statistical numeracy and risk literacy of participants (Cokely et al., 2012). It is a four-question, fill-in-the-blank assessment. The Berlin Numeracy Test has been validated as a measure of numeracy in highly-educated samples in multiple studies in the U.S. and internationally (Ghazal et al., 2014; Petrova, Van der Pligt, & Garcia-Retamero, 2014). The test had a Cronbach’s alpha of .84 in a sample of 294 surgeons from 60 countries (Garcia-Retamero et al., 2014). The test error/reliability of a brief instrument such as this is best measured by test-retest reliability rather than Cronbach’s alpha. The developers of the Berlin Numeracy Test reported test-retest reliability of r > .98 (Cokely et al., 2012).

Cokely et al. (2012, p. 26) explained three methods used to measure statistical numeracy and risk literacy with the most common form being “the use of objective performance measures of numeracy” such as in psychometric tests. In contrast, subjective measures of numeracy focused on clinical and health domains have been developed. Additionally, the relationship between risky decisions and “overall educational attainment, cognitive abilities, or cognitive styles” is often studied. The developers of the Berlin Numeracy Test sought to build upon previous measures of numeracy to develop a brief instrument with improved discriminability for highly-educated samples. Cokely et al. (2012, p. 26) claimed to design a “fast, valid, and reliable tool for research, assessment, and public outreach.” The test is estimated to take less than four minutes. Permission to use the instrument was obtained via email from Dr. Edward Cokely (see
The researcher scored the Berlin Numeracy Test by checking participant responses against the key provided by the instrument developers. Each question is given a 1-point value giving the test a possible score of 0 to 4 in whole number increments. The developers do not assign meaning to specific scores, but 0 indicates no evidence of risk literacy while a 4 indicates maximum risk literacy. Score on the Berlin Numeracy Test have typically been used in correlational analyses with other measures (Cokely et al., 2012; Garcia-Retamero et al., 2014; Ghazal et al., 2014).

The ATS scale (see Appendix C for instrument) was used to measure the predictor variables, attitude toward the field of statistics and attitude toward the course of statistics (Schultz & Koshino, 1998; Wise, 1985). The ATS is a 29-question Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The assessment is comprised of two subscales: the Attitude Toward the Field of Statistics subscale (20 items) and the Attitude Toward the Course of Statistics subscale (9 items). The ATS scale and the individual subscales have been validated in multiple studies (Cashin & Elmore, 2005; Kottke, 2000; Perepiczka, Chandler, & Becerra, 2011; Rhoads & Hubele, 2000; Roberts & Reese, 1987; Vanhoof, 2006; Waters, Martelli, Zakrajsek, & Popovich, 1988; Wise, 1985). The entire 29-item scale had a Cronbach’s alpha ranging from 0.89 to 0.94 in four studies (Cashin & Elmore, 2005; Kottke, 2000; Roberts & Reese, 1987; Vanhoof, 2006). The 20-item Attitude Toward the Field of Statistics subscale had a Cronbach’s alpha ranging from 0.83 to 0.96 in seven studies (Cashin & Elmore, 2005; Kottke, 2000; Perepiczka et al., 2011; Rhoads & Hubele, 2000; Vanhoof, 2006; Waters et al., 1988; Wise, 1985). The 9-item Attitude Toward the Course of Statistics subscale had a Cronbach’s alpha ranging from 0.77 to 0.92 in seven studies (Cashin & Elmore, 2005; Kottke, 2000;
Perepiczka et al., 2011; Rhoads & Hubele, 2000; Vanhoof, 2006; Waters et al., 1988; Wise, 1985). The developer of the scale, Dr. Stephen Wise, has given permission for researchers to use the instrument (see Appendix D).

Wise (1985) described limitations of an instrument used for measuring student attitudes toward statistics prior to the development of the ATS scale. The 34-item Statistics Attitude Survey (SAS) was often used by instructors pre- and post-statistics courses with scores highly correlated to course grades. However, Wise posited that one third or more of the tool’s questions measured student achievement rather than attitudes. Furthermore, Wise suggested many of the SAS questions were not appropriate for students enrolled in an introductory statistics course who lacked the knowledge needed to answer the questions. The ATS scale was constructed as an alternative scale to specifically measure student attitudes without the limitations of the SAS.

The researcher scored the ATS scale by totaling the responses on the Likert scales for each subscale. Each question has a value ranging from 1 (strongly disagree) to 5 (strongly agree). The Attitude Toward the Field of Statistics subscale includes questions 1, 2, 5, 6*, 9, 10*, 11, 13, 14*, 16*, 17, 19, 20*, 21, 22, 23, 24, 26, 28*, and 29 (* represents reverse coded) resulting in an interval value ranging from 20 to 100 in whole number increments. The Attitude Toward the Course of Statistics subscale includes questions 2*, 4*, 7*, 8, 12*, 15*, 18*, 25*, 27* resulting in an interval value ranging from 9 to 45. Additionally, the overall ATS scale score was calculated by summing the subscales with the possible range in score of 29 to 145.

The Berlin Numeracy Test and ATS scale were part of a questionnaire that also included the purpose of the study, consent forms (no signature required), instructions, demographic questions (age, gender, ethnicity, race), and a question of “Did you successfully complete a college level statistics course with a grade of C or better. The total time to complete the entire
questionnaire was less than 15 minutes. Responses were graded by the researcher and recorded in SPSS.

**Procedures**

Institutional Review Board (IRB) approval was obtained from Liberty University and the two institutions where the research took place. First, exempt IRB status was obtained from Liberty University prior to applying for approval at the colleges of medicine. University A did not allow Liberty IRB oversight but also determined the research as exempt. University B did allow Liberty IRB oversight and determined the research as exempt. A paper, hard copy questionnaire was created to include explanation of the purpose of the study, consent forms (no signature required), instructions, demographic questions (age, gender, ethnicity, race), question of “Did you successfully complete a college level statistics course with a grade of C or better?” the 29-item ATS scale, and the four questions for the Berlin Numeracy Test. The researcher explained the purpose and protocol of the study to the Associate Dean at University A (see Appendix E) and the Chair of the Curriculum Committee at University B (see Appendix F) and requested they allow administration of the survey to students during a class by a researcher affiliated with the study. The researcher explained that participant names would be put in drawing for a $100 Amazon Gift Card at each university if participants noted in their survey they want to be included and provided an email address.

The researcher at University A and a co-researcher at University B administered and collected the surveys during a class. The last page of the survey included the option of providing an email address to be entered into the drawing for the gift card. The page was removed immediately following the administration of the questionnaire, and a winner was randomly selected. The gift card was given directly to the winning student at University A, and an e-card
was sent to the winning student at University B. The co-researcher from University B stored
surveys in a locked filing cabinet prior to sending them to the investigator via certified mail. The
researcher at University A stored surveys from both institutions in a locked filing cabinet,
although no personal identifying information remained on the surveys. The researcher scored the
assessments and recorded data in SPSS analytical software. Appropriate analysis was conducted
and results are reported.

**Data Analysis**

Warner (2013) recommends using a regression analysis in non-experimental research in
which the researcher has not manipulated the variables and causal inferences cannot be made.
Multiple regression is suggested when the effect of more than two predictors on one criterion
variable is being tested. This quantitative, non-experimental, predictive, correlational study
examined the relationship between medical students’ risk literacy as measured by the Berlin
Numeracy Test and a linear combination of scores on The ATS subscales and successful
completion of a college level statistics course.

A standard, or simultaneous, multiple regression was performed. Contrary to sequential
or hierarchical method of entry or statistical order of entry, with this analysis, the predictor
variables are all entered at once with coefficients calculated for one regression equation
representing all predictors (Warner, 2013). Standard or simultaneous data entry allows for the
most conservative assessment of the prediction of individual variable effect while controlling for
the other predictors. Typically, with standard regressions, the proportion of variance in the
criterion variable that is due to a specific predictor variable is less than when the predictor
variable is entered in a hierarchal or statistical entry method. Pairwise deletion was used to
maximize Ns for all variables evaluated. Linear regression was also performed to test the each of
the additional hypotheses relating the criterion variable to individual predictor variables.

Assumptions for multiple regression include the variables should be measured on the interval or ratio level, but the predictor variable may be categorical (Warner, 2013). The criterion variable, score on the Berlin Numeracy Test, is measured on the ratio level. The predictor variables, scores on the ATS subscales, are measured on the interval level. The predictor variable, successful completion of a college level statistics course, is a dichotomous, categorical variable, yes or no, and was coded 1 for yes and 0 for no in SPSS (Green & Salkind, 2013). The observations within each variable were independent. This study assumed that all predictor variables and the criterion variable were measured reliably and without error. The instruments used in the study have established reliability and validity; however, error could be introduced if the students made a mistake entering information into the survey.

Initial data screening included checking for outliers using a histogram for each interval predictor variable and the criterion variable (Warner, 2013). Quantitative variables must be normally distributed, especially the criterion variable, and extreme outliers removed. Removing outliers can reduce the probability of both Type I and Type II error (Osborne & Waters, 2002). Scatter plots were made for each pair of interval variables (score on Attitude Toward the Field of Statistics subscale with score on Berlin Numeracy Test and score on Attitude Toward the Course of Statistics subscale with score on Berlin Numeracy Test). The scatterplots were used to check for outliers but also to check for the assumption of linearity; the outcome variable should be linearly related to the predictor variable. The categorical variable of successful completion of a college level statistics course was coded 0 and 1 for SPSS analysis but was not included in these initial data screening tests since it is a dichotomous variable (Green & Salkind, 2013).

Further assumption tests for linearity included plots of the standardized residuals against
the predicted standardized residuals for each regression model. Observation of a random distribution with no obvious pattern of positive and negative values would support linearity. The same plots were used to test the assumption of independent errors, for any pair of observations, the error terms should be uncorrelated (Myers & Well, 2003). Successive residuals should be independent with no pattern or high correlations; no long runs of positive or negative residuals should be observed in the plots to confirm independent errors.

Additionally, multiple regression analysis has an assumption of homoscedasticity that requires error variance to be equal across all levels of the predictor variables (Warner, 2013). This requires the variation about the predicted values is constant regardless whether the predicted values are large or small. The same plots of standardized residuals against the predicted standardized residuals used to test the assumptions of linearity and independent errors were used to check for homoscedasticity.

Regression analysis assumes no multicollinearity among independent variables (Myers & Well, 2003). When two or more variables are closely linearly related, the standard error of beta coefficients in a multiple regression increases. Bivariate correlations among the independent variables were examined to check for multicollinearity using the values of the Pearson correlation coefficient and the variance inflation factor. This assumption test requires \(-0.70 \leq r \leq 0.70\) and \(\text{VIF} < 10\) for each set of variables.

Finally, the assumption of normally distributed errors was tested by again using the plots of standardized residuals against the predicted standardized residuals, normal probability plots, and histograms for each regression model. The normal probability plots should reveal generally a diagonal line for the observed cumulative probability plotted against the expected cumulative probability. The histograms for the regression standardized residuals should generally have a
normal distribution.

Descriptive statistics were explored, mean (\(M\)) and standard deviation (\(SD\)), for the interval level predictor variables and the criterion variable. The correlation between each individual predictor variable and the criterion variable were explored to determine independent contributions without considering or controlling for the other variables. Multiple correlation value (\(R\)), squared multiple correlation value (\(R^2\)), adjusted squared multiple correlation value (\(R^2_{adj}\)), \(F\) value (\(F\)), and significance level (\(p\)) are reported. The effect of individual predictor variables was assessed by examining the individual regression slope \(t\) ratios. Those values determined both the strength and direction of the predictor variable contributions. These results determined if the null hypotheses are rejected or fail to be rejected.
CHAPTER FOUR: FINDINGS

Overview

This chapter presents the results of the data analysis starting with descriptive statistics followed by multiple regression analysis findings that address the research question and null hypotheses. The descriptive statistics include data such as the mean and standard deviation of groups on the criterion variable, the Berlin Numeracy Test, as well as the predictor variables, the Attitude Toward the Field of Statistics and the Attitude Toward the Course of Statistics. Groups are reported by study site and demographics. Further details are provided for overall sample answers on the Berlin Numeracy Test and the two ATS subscales. Finally, the responses to the question, “Did you successfully complete a college level statistics course with a grade of C or better?” is reported by data collection site. Assumption tests are reported with figures and tables prior to detailed statistics and evaluation of each null hypothesis.

Research Question(s)

RQ1: How accurately can risk literacy as measured by the Berlin Numeracy Test be predicted from a linear combination of score on the Attitude Toward the Field of Statistics subscale, score on the Attitude Toward the Course of Statistics subscale, and successful completion of a college level statistics course?

Null Hypotheses

H₀₁: There is no statistically significant predictive relationship between risk literacy as measured by the Berlin Numeracy Test and the linear combination of score on the Attitude Toward the Field of Statistics subscale, score on the Attitude the Toward Course of Statistics subscale, and successful completion of a college level statistics course.
**H₀2:** There is no statistically significant predictive relationship between risk literacy as measured by the Berlin Numeracy Test and score on the Attitude Toward the Field of Statistics subscale.

**H₀3:** There is no statistically significant predictive relationship between risk literacy as measured by the Berlin Numeracy Test and score on the Attitude Toward the Course of Statistics subscale.

**H₀4:** There is no statistically significant predictive relationship between risk literacy as measured by the Berlin Numeracy Test and successful completion of a college level statistics course.

**Descriptive Statistics**

Analysis of the surveys revealed the mean for the criterion variable, score on the Berlin Numeracy Test, the measurement for risk literacy, was 2.06 for the overall sample. The minimum and maximum scores possible were 0 and 4, and the sample had the same minimum and maximum scores. The mean scores for two of the predictor variables, the Attitude Toward the Field of Statistics subscale and the Attitude Toward the Course of Statistics subscale, were 78.60 and 32.54, respectively. The range of scores possible on the two subscales was 20-100 and 9-45, respectively. Table 2 displays these overall sample results as well as individual university means. University B outscored University A for all variables; however, the differences were only statistically significant for the Attitude Toward the Field of Statistics subscale ($p = .000$).

Further breakdown by gender, ethnicity, and race revealed differences in risk literacy and attitudes toward statistics both on the field and course subscales (see Table 3). Males (n = 144) had a mean score of 2.28 on the Berlin Numeracy Test while their female counterparts (n = 183) had a mean score of 1.89, a statistically significant difference ($p = .004$). Hispanic participants
Table 2.

*Means and Standard Deviations of Variables by University and Overall*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Demographic</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin Numeracy Test</td>
<td>University A</td>
<td>172</td>
<td>1.95</td>
<td>1.16</td>
</tr>
<tr>
<td></td>
<td>University B</td>
<td>155</td>
<td>2.19</td>
<td>1.27</td>
</tr>
<tr>
<td></td>
<td>Total Sample</td>
<td>327</td>
<td>2.06</td>
<td>1.21</td>
</tr>
<tr>
<td>Attitude Toward the Field of Statistics</td>
<td>University A</td>
<td>172</td>
<td>75.91</td>
<td>10.34</td>
</tr>
<tr>
<td></td>
<td>University B</td>
<td>155</td>
<td>81.59</td>
<td>8.44</td>
</tr>
<tr>
<td></td>
<td>Total Sample</td>
<td>327</td>
<td>78.60</td>
<td>9.99</td>
</tr>
<tr>
<td>Attitude Toward the Course of Statistics</td>
<td>University A</td>
<td>172</td>
<td>31.95</td>
<td>6.32</td>
</tr>
<tr>
<td></td>
<td>University B</td>
<td>155</td>
<td>33.20</td>
<td>6.33</td>
</tr>
<tr>
<td></td>
<td>Total Sample</td>
<td>327</td>
<td>32.54</td>
<td>6.35</td>
</tr>
</tbody>
</table>

*Note.* Possible scores range from 0 to 4, 20-100, and 9-45 on the Berlin Numeracy Test, Attitude Toward the Field of Statistics, and Attitude Toward the Course of Statistics, respectively.
### Table 3

**Means and Standard Deviations of Variables by Demographic**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Demographic</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Berlin Numeracy Test</strong></td>
<td>Male</td>
<td>144</td>
<td>2.28</td>
<td>1.20</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>183</td>
<td>1.89</td>
<td>1.20</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>0</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td></td>
<td>Hispanic</td>
<td>23</td>
<td>1.43</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>Not Hispanic</td>
<td>303</td>
<td>2.11</td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>Black/African American</td>
<td>10</td>
<td>.70</td>
<td>.95</td>
</tr>
<tr>
<td></td>
<td>American Indian/Native Alaskan</td>
<td>3</td>
<td>2.77</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>Native Hawaiian/Pacific Islander</td>
<td>3</td>
<td>1.33</td>
<td>.58</td>
</tr>
<tr>
<td></td>
<td>White</td>
<td>263</td>
<td>2.16</td>
<td>1.16</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>42</td>
<td>1.88</td>
<td>1.37</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>14</td>
<td>1.21</td>
<td>1.25</td>
</tr>
<tr>
<td><strong>Attitude Toward the Field of Statistics</strong></td>
<td>Male</td>
<td>144</td>
<td>79.22</td>
<td>9.17</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>183</td>
<td>78.11</td>
<td>10.41</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>0</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td></td>
<td>Hispanic</td>
<td>23</td>
<td>78.35</td>
<td>10.9</td>
</tr>
<tr>
<td></td>
<td>Not Hispanic</td>
<td>303</td>
<td>78.68</td>
<td>9.80</td>
</tr>
<tr>
<td></td>
<td>Black/African American</td>
<td>10</td>
<td>72.70</td>
<td>10.47</td>
</tr>
<tr>
<td></td>
<td>American Indian/Native Alaskan</td>
<td>3</td>
<td>84.00</td>
<td>13.08</td>
</tr>
<tr>
<td></td>
<td>Native Hawaiian/Pacific Islander</td>
<td>3</td>
<td>74.67</td>
<td>10.69</td>
</tr>
<tr>
<td></td>
<td>White</td>
<td>263</td>
<td>78.79</td>
<td>9.51</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>42</td>
<td>80.38</td>
<td>10.86</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>14</td>
<td>79.14</td>
<td>12.02</td>
</tr>
<tr>
<td><strong>Attitude Toward the Course of Statistics</strong></td>
<td>Male</td>
<td>144</td>
<td>32.85</td>
<td>5.85</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>183</td>
<td>32.30</td>
<td>6.72</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>0</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td></td>
<td>Hispanic</td>
<td>23</td>
<td>31.65</td>
<td>7.13</td>
</tr>
<tr>
<td></td>
<td>Not Hispanic</td>
<td>303</td>
<td>32.64</td>
<td>6.29</td>
</tr>
<tr>
<td></td>
<td>Black/African American</td>
<td>10</td>
<td>30.50</td>
<td>8.72</td>
</tr>
<tr>
<td></td>
<td>American Indian/Native Alaskan</td>
<td>3</td>
<td>35.67</td>
<td>5.03</td>
</tr>
<tr>
<td></td>
<td>Native Hawaiian/Pacific Islander</td>
<td>3</td>
<td>30.00</td>
<td>7.55</td>
</tr>
<tr>
<td></td>
<td>White</td>
<td>263</td>
<td>32.61</td>
<td>6.16</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>42</td>
<td>33.29</td>
<td>6.50</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>14</td>
<td>32.43</td>
<td>8.35</td>
</tr>
</tbody>
</table>
(n = 23) had statistically significant lower scores \( (p = .010) \) on the Berlin Numeracy Test than their non-Hispanic counterparts \( (n = 303) \) with means of 1.43 and 2.11, respectively. The mean score of Black/African American participants on the Berlin Numeracy Test was 0.70 while non-Black/African American participants had a mean score of 2.11, also a statistically significant difference \( (p = .000) \). However, there were only 10 participants who self-reported as black, and participants were allowed to select more than one racial category. Differences in attitudes by gender, ethnicity, and race did not reach statistical significance except for Black/African American participants had lower scores on the Attitude Toward the Field of Statistics subscale \( (p = .049) \).

Analysis of the individual questions on the Berlin Numeracy Test revealed the majority of participants correctly answered questions 1 and 2 but incorrectly answered questions 3 and 4 (see Figure 4). The first question was answered correctly by about three-fourths of respondents while the last question was answered incorrectly by nearly the same portion. Total scores on the Berlin Numeracy Test are displayed in Figure 5. Examining participant performance by study site revealed 28.5% of University A participants scored at the higher level of a 3 or 4 while 41.9% of University B participants scored in that range. Individual question responses to the Attitude Toward the Field of Statistics and Attitude Toward the Course of Statistics subscales are displayed in Tables 4 and 5, respectively.
Figure 4. Percentage correct and incorrect by question on the Berlin Numeracy Test.

Figure 5. Percentages of scores on the Berlin Numeracy Test.
### Table 4

**Attitude Toward the Field of Statistics Subscale Response Percentages**

<table>
<thead>
<tr>
<th>Question</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I feel that statistics will be useful to me in my profession.</td>
<td>.3</td>
<td>2.8</td>
<td>10.4</td>
<td>54.7</td>
<td>31.8</td>
</tr>
<tr>
<td>3. A good researcher must have training in statistics.</td>
<td>.3</td>
<td>1.2</td>
<td>6.7</td>
<td>45.0</td>
<td>46.8</td>
</tr>
<tr>
<td>5. Most people would benefit from taking a statistics course.</td>
<td>.6</td>
<td>1.8</td>
<td>20.2</td>
<td>56.9</td>
<td>20.5</td>
</tr>
<tr>
<td>6. I have difficulty seeing how statistics relates to my field of study.*</td>
<td>35.5</td>
<td>52.6</td>
<td>8.0</td>
<td>3.4</td>
<td>.6</td>
</tr>
<tr>
<td>9. Statistics will be useful to me in comparing the relative merits of different objects, methods, programs, etc.</td>
<td>.3</td>
<td>3.1</td>
<td>14.4</td>
<td>58.1</td>
<td>24.4</td>
</tr>
<tr>
<td>10. Statistics is not really very useful because it tells us what we already know anyway.*</td>
<td>43.7</td>
<td>50.8</td>
<td>3.7</td>
<td>1.5</td>
<td>.3</td>
</tr>
<tr>
<td>11. Statistical training is relevant to my performance in my field of study.</td>
<td>.6</td>
<td>8.3</td>
<td>23.2</td>
<td>51.4</td>
<td>16.5</td>
</tr>
<tr>
<td>13. Statistics is a worthwhile part of my professional training.</td>
<td>.6</td>
<td>5.2</td>
<td>20.8</td>
<td>57.2</td>
<td>16.2</td>
</tr>
<tr>
<td>14. Statistics is too math oriented to be of much use to me in the future.*</td>
<td>32.4</td>
<td>56.0</td>
<td>9.2</td>
<td>2.1</td>
<td>.3</td>
</tr>
<tr>
<td>16. Statistical analysis is best left to the &quot;experts&quot; and should not be part of a lay professional's job.*</td>
<td>15.0</td>
<td>58.1</td>
<td>19.3</td>
<td>7.3</td>
<td>.3</td>
</tr>
<tr>
<td>17. Statistics is an inseparable aspect of scientific research.</td>
<td>0</td>
<td>1.2</td>
<td>4.0</td>
<td>54.1</td>
<td>40.7</td>
</tr>
</tbody>
</table>

*Reverse keyed items
<table>
<thead>
<tr>
<th>Question</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>19. I am excited at the prospect of actually using statistics in my job.</td>
<td>7.0</td>
<td>30.0</td>
<td>39.4</td>
<td>17.4</td>
<td>6.1</td>
</tr>
<tr>
<td>20. Studying statistics is a waste of time.*</td>
<td>40.1</td>
<td>50.8</td>
<td>7.3</td>
<td>.9</td>
<td>.9</td>
</tr>
<tr>
<td>21. My statistical training will help me better understand the research being done in my field of study.</td>
<td>.3</td>
<td>1.8</td>
<td>8.6</td>
<td>59.3</td>
<td>30.0</td>
</tr>
<tr>
<td>22. One becomes a more effective &quot;consumer&quot; of research findings if one has some training in statistics.</td>
<td>.3</td>
<td>1.5</td>
<td>5.5</td>
<td>56.6</td>
<td>36.1</td>
</tr>
<tr>
<td>23. Training in statistics makes for a more well-rounded professional experience.</td>
<td>0</td>
<td>1.2</td>
<td>10.4</td>
<td>63.9</td>
<td>24.5</td>
</tr>
<tr>
<td>24. Statistical thinking can play a useful role in everyday life.</td>
<td>.3</td>
<td>5.2</td>
<td>20.2</td>
<td>54.7</td>
<td>19.6</td>
</tr>
<tr>
<td>26. I feel that statistics should be required early in one's professional training.</td>
<td>.6</td>
<td>5.2</td>
<td>28.7</td>
<td>55.4</td>
<td>10.1</td>
</tr>
<tr>
<td>28. Statistical training is not really useful for most professionals.*</td>
<td>20.2</td>
<td>54.4</td>
<td>21.1</td>
<td>4.0</td>
<td>.3</td>
</tr>
<tr>
<td>29. Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.</td>
<td>14.4</td>
<td>40.1</td>
<td>30.6</td>
<td>11.9</td>
<td>3.1</td>
</tr>
</tbody>
</table>

*Reverse keyed items
Table 5

*Reverse keyed items
A primary expected difference between data collection sites was the participant responses to the survey question, “Did you successfully complete a college level statistics course with a grade of C or better?” University A requires a statistics course as a prerequisite while University B does not; however, the difference was minimal with 96.5% of University A students answering “yes” to the question and 89.0% of University B students (see Figure 5). This resulted in disproportionate groups for the predictor variable for this measure.

![Figure 6](image)

*Figure 6.* Responses to the question, “Did you successfully complete a college level statistics course with a grade of C or better?”
Results

Hypotheses

Assumption Tests

All hypotheses used similar statistical analysis, linear regression models; therefore, the same assumption tests and data screening were conducted for each prior to running the regression analyses. As stated in the data analysis section, the criterion variable, score on the Berlin Numeracy Test, is measured on the ratio level. The predictor variables, scores on the ATS subscales, are measured on the interval level. The predictor variable, successful completion of a college level statistics course, is a dichotomous, categorical variable, yes or no, and was coded 0 and 1 in SPSS. The following assumptions were tested: (a) no significant outliers, (b) linear relationship between the predictor variables and the criterion variable, (c) linearity for residuals, (d) homoscedasticity, (e) absence of multicollinearity, and (f) independent errors. The figures displaying the graphs for these assumption tests follow.

Histograms for each interval predictor variable and the criterion variable revealed no extreme outliers (see Figures 7, 8, and 9). Scatter plots for each pair of interval variables further revealed no extreme outliers and confirmed linear relationship between the outcome variable and each predictor variable (see Figures 10 and 11). Plots of the standardized residuals against the predicted standardized residuals confirmed linearity for each regression model (see Figures 12, 13, and 14). Random distributions with no obvious pattern of positive and negative values were observed. Additionally, the plots of standardized residuals against the predicted standardized residuals confirmed homoscedasticity, equal error variance across all levels of the predictor variables. This variation about the predicted values was constant regardless if the predicted values were large or small.
Figure 7. Histogram of Berlin Numeracy Test scores.

Figure 8. Histogram of Attitude Toward the Field of Statistics scores.
Figure 9. Histogram of Attitude Toward the Course of Statistics scores.

Bivariate correlations among the independent variables were examined to check for multicollinearity using the values of the Pearson correlation coefficient and the variance inflation factor. The Pearson correlation coefficients were all within range; however there was significant correlation between the ATS subscales (see Table 6). The VIF for each regression model also fulfilled this assumption test requirement (see Table 7). The plots of standardized residuals against the predicted standardized residuals were used along with normal probability plots (see Figures 15, 16, 17, and 18) and histograms (see Figures 19, 20, 21, and 22) for each regression model to confirm the assumption of normally distributed errors. Successive residuals were independent with no pattern or high correlations; no long runs of positive or negative residuals were observed. The normal probability plots revealed generally a diagonal line for the observed cumulative probability plotted against the expected cumulative probability. The histograms for the regression standardized residuals generally had a normal distribution.
Figure 10. Scatterplot of Attitude Toward the Field of Statistics and Berlin Numeracy Test scores.

Figure 11. Scatterplot of Attitude Toward the Course of Statistics and Berlin Numeracy Test scores.
Figure 12. Plots of the standardized residuals and the predicted standardized residuals for the regression model with the predictor variable Attitude Toward the Field of Statistics and criterion variable Berlin Numeracy Test.

Figure 13. Plots of the standardized residuals and the predicted standardized residuals for the regression model with the predictor variable Attitude Toward the Course of Statistics and criterion variable Berlin Numeracy Test.
Figure 14. Plots of the standardized residuals and the predicted standardized residuals for the regression model with the predictor variable Successful Completion of a College Level Statistics Course and criterion variable Berlin Numeracy Test.
Table 6

*Pearson Correlation Coefficients of Independent Variables*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Attitude Toward the Field of Statistics</th>
<th>Attitude Toward the Course of Statistics</th>
<th>Successful Completion of College Level Statistics Course</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitude Toward the Field of Statistics</td>
<td>1</td>
<td>.593**</td>
<td>.000</td>
</tr>
<tr>
<td>Attitude Toward the Course of Statistics</td>
<td>.593**</td>
<td>1</td>
<td>-.007</td>
</tr>
<tr>
<td>Successful Completion of College Level Statistics Course</td>
<td>.000</td>
<td>-.007</td>
<td>1</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

Table 7

*Variance Inflation Factors*

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Attitude Toward the Field of Statistics</th>
<th>Attitude Toward the Course of Statistics</th>
<th>Successful Completion of College Level Statistics Course</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothesis 1</td>
<td>1.543</td>
<td>1.543</td>
<td>1.000</td>
</tr>
<tr>
<td>Hypotheses 2</td>
<td>1.000</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Hypothesis 3</td>
<td>.</td>
<td>1.000</td>
<td>.</td>
</tr>
<tr>
<td>Hypothesis 4</td>
<td>.</td>
<td>.</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Figure 15. Normal probability plot of regression standardized residuals for Hypothesis 1.

Figure 16. Normal probability plot of regression standardized residuals for Hypothesis 2.
Figure 17. Normal probability plot of regression standardized residuals for Hypothesis 3.

Figure 18. Normal probability plot of regression standardized residuals for Hypothesis 4.
Figure 19. Histogram of regression standardized residuals for Hypothesis 1.

Figure 20. Histogram of regression standardized residuals for Hypothesis 2.
Figure 21. Histogram of regression standardized residuals for Hypothesis 3.

Figure 22. Histogram of regression standardized residuals for Hypothesis 4.
Null Hypotheses

**H01**: There is no statistically significant predictive relationship between risk literacy as measured by the Berlin Numeracy Test and the linear combination of score on the Attitude Toward the Field of Statistics subscale, score on the Attitude the Toward Course of Statistics subscale, and successful completion of a college level statistics course.

Each predictor variable, score on the Attitude Toward the Field of Statistics subscale, score on the Attitude the Toward Course of Statistics subscale, and successful completion of a college level statistics course, were entered as independent variables with the criterion variable entered as the predictor variable in a multiple regression analysis in SPSS. A significant regression equation was found ($F(3, 323) = 3.109, p = .027$), with $R = .168$, $R^2 = .028$, and $R^2_{adj} = .019$. Individual t tests for the coefficients in this model were $t(323) = 1.248, p = .213$; $t(323) = 1.131, p = .259$; and $t(323) = -1.529, p = .127$ for $x_1$, $x_2$, and $x_3$, respectively where $x_1$ is the score on the Attitude Toward the Field of Statistics subscale, $x_2$ is the Attitude Toward the Course of Statistics subscale, and $x_3$ is the variable coded yes = 1 and no = 0 for successful completion of a college level statistics course. The regression equation for predicting participants’ Berlin Numeracy Test score was

$$y = 1.127 + .010x_1 + .015x_2 - .399x_3.$$ 

Participants’ Berlin Numeracy Test score increased .010 for each point on the Attitude Toward the Field of Statistics subscale and .015 for each point on the Attitude Toward the Course of Statistics subscale. Participants that successfully completed a college level statistics course scored .399 lower on the Berlin Numeracy Test than participants who did not. The results suggested rejection of the null hypothesis.
**H02:** There is no statistically significant predictive relationship between risk literacy as measured by the Berlin Numeracy Test and score on the Attitude Toward the Field of Statistics subscale.

The correlation between the score on the Berlin Numeracy Test and the Attitude Toward the Field of Statistics when examined in a linear relationship without considering the other predictor variables, had a statistically significant predictive relationship ($F(1, 325) = 5.652$, $p = .018$), with $R = .131$, $R^2 = .017$, $R^2_{adj} = .014$, and $t(325) = 2.377$. The corresponding regression equation was

$$y = .794 + .016x,$$

where $y$ represented the score on the Berlin Numeracy Test and $x$ represented the total on the Attitude Toward the Field of Statistics subscale. The results supported rejection of the null hypothesis.

**H03:** There is no statistically significant predictive relationship between risk literacy as measured by the Berlin Numeracy Test and score on the Attitude Toward the Course of Statistics subscale.

The correlation between the score on the Berlin Numeracy Test and the Attitude Toward the Course of Statistics when examined in a linear relationship without considering the other predictor variables, had a statistically significant predictive relationship ($F(1, 325) = 5.417$, $p = .021$), with $R = .128$, $R^2 = .016$, $R^2_{adj} = .013$, and $t(325) = 2.327$. The corresponding regression equation was

$$y = 1.261 + .025x,$$

where $y$ represented the score on the Berlin Numeracy Test and $x$ represented the total on the Attitude Toward the Course of Statistics subscale. The results supported rejection of the null hypothesis.
hypothesis.

**H0**: There is no statistically significant predictive relationship between risk literacy as measured by the Berlin Numeracy Test and successful completion of a college level statistics course.

The correlation between the score on the Berlin Numeracy Test and the successful completion of a college level statistics course when examined in a linear relationship without considering the other predictor variables, did not have a statistically significant predictive relationship (F(1, 325) = 2.332, \( p = .128 \)), with \( R = .084, R^2 = .007, R^2_{adj} = .004 \), and \( t(325) = -1.527 \). The results did not support rejection of the null hypothesis.
CHAPTER FIVE: CONCLUSIONS

Overview

This chapter begins with a discussion of the findings of this study and comparison with the related literature. Interpretation of results for the data analysis of each variable are reviewed as well as the specific outcomes pertaining to the research question and four hypotheses. The implications of this study for medical school curriculum on risk literacy and statistical numeracy are suggested, and limitations that could pose threat to the validity or generalizability of the study are considered. Recommendations for future research on medical school curriculum for health numeracy topics including skills needed to effectively communicate risk are proposed.

Discussion

The purpose of this quantitative, non-experimental, predictive, correlational study was to examine the relationship between medical students’ risk literacy as measured by the Berlin Numeracy Test and a linear combination of the Attitude Toward the Field of Statistics subscale, the Attitude Toward the Course of Statistics subscale, and successful completion of a college level statistics course. Findings of this study regarding medical students’ risk literacy aptitude were comparable to others in the literature with 34.9% of the sample scoring in the upper two categories on the Berlin Numeracy Test, 55.1% in the bottom two categories, and 10% missing all four questions. Similarly, in an international sample of surgeons from 60 countries, 50% scored in the bottom two categories on the Berlin Numeracy Test (Garcia-Retamero et al., 2014). Studies have consistently shown that the medical community and those training as health professionals lack adequate statistical knowledge on assessments ranging from those that focus on risk literacy like the Berlin Numeracy Test and others that concentrate specifically on biostatistics (Anderson et al., 2014; Berwick et al., 1981; Johnson et al., 2014; Susarla & Redett,
The mean of the Attitude Toward the Field of Statistics subscale was 78.60 on a scale of 20-100 possible, and the mean of the Attitude Toward the Course of Statistics subscale was 32.54 on a scale of 9-45 possible. Participants had a more favorable attitude toward the course of statistics than toward the field, but overall the mean attitude for either domain was more positive than negative on the Likert scale rating. Measures for attitudes towards statistics had inconsistent outcomes in the literature with the majority of studies reporting negative attitudes of health profession students toward statistics in general and often not seeing the need for statistics in their career, also referred to as utility (Beurze et al., 2013; Bookstaver et al., 2012; Hannigan et al., 2014; Kiekkas et al., 2015; Zhang, 2012). However, some studies have shown positive attitudes toward different domains of statistics (Baghi & Kornides, 2013). Other studies included additional constructs such as self-confidence and anxiety toward statistics.

It is known that the vast majority of medical schools do not require a college level statistics course as a prerequisite (see Table 1), but it was not reported what portion of medical students enter their training having successfully completed such a course. This study expected that all students from University A would have successfully completed a college level statistics course since it is a prematriculation requirement and that a much smaller portion of students from University B would have since it is not required by that institution. However, this study found 3.5% of University A and 11.0% of University B students reported having not successfully completed a college level statistics course.

**RQ1:** How accurately can risk literacy as measured by the Berlin Numeracy Test be predicted from a linear combination of score on the Attitude Toward the Field of Statistics
subscale, score on the Attitude Toward the Course of Statistics subscale, and successful completion of a college level statistics course?

A multiple regression was used to evaluate the research question as well as linear regressions to examine the relationship between individual predictor variables and the criterion variable. Although a statistically significant predictive relationship was found, it was a moderate correlation, and the direction of the relationship for successful completion of a college level statistics course was counter-intuitive. As expected, both domains for attitudes toward statistics had positive correlations with the outcome variable, score on the Berlin Numeracy Test when examined alone in linear regressions or in the combination regression model. The individual contributions of the ATS subscales were not statistically significant in the multiple regression but were statistically significant when evaluated as single predictors in linear models. These results were consistent with other studies that have examined correlations among attitudes toward statistics and math knowledge or performance (Bookstaver et al., 2012; Zhang et al., 2012). The instruments from this study have not previously been used together to compare the constructs of risk literacy, attitude toward the field of statistics, and attitude toward the course of statistics. The Berlin Numeracy Test measures risk literacy, which focuses on probability with fractions and percentages while some assessments in the literature focused on broad mathematical concepts or final summative exams in a statistics course (Cokely et al., 2012; Kiekkas et al., 2015; Thompson et al., 2015). Also, participants in this study were not allowed to use calculators for computations while other studies did not report this constraint when measuring statistical knowledge.

The variable successful completion of a college level statistics course was expected to positively predict score on the Berlin Numeracy Test, but the multiple regression model and the
individual linear regression model both showed a negative relationship. This variable contributed to an overall significant predictive relationship for the linear combination of predictor variables, but results were not significant for the individual variable alone. The unexpected results may be attributed to the unequal groups for this variable, as previously discussed, with over 90% of the sample, 304 out of 327 participants, reporting having successfully completed a college level statistics course. Similar comparisons in the literature supported statistics coursework improving both attitudes toward statistics and performance, but none of the studies compared risk literacy and completion of a college level statistics course (Baghi & Kornides, 2013; Kiekkas et al., 2015).

The results of this study supported previous findings that many medical students lack adequate statistical knowledge and that attitudes towards statistics in multiple domains are positively correlated with statistical knowledge. This added to the literature by focusing on risk literacy, which is specifically needed by medical practitioners to understand and communicate risk, rather than advanced statistics or biostatistics that was often the focus in the literature (Ghazal et al., 2014; Johnson et al., 2014; Trickey et al., 2014). Previous studies have not adequately explored the impact of prerequisite statistics courses on risk literacy, and this study did not add valuable findings about that relationship.

**Implications**

Health numeracy is a complex concept that requires both patients and healthcare providers to “access, process, interpret, communicate and act on numerical, quantitative, graphical, biostatistical, and probabilistic health information…to make effective health decisions” (Golbeck, Ahlers-Schmidt, Paschal, & Dismuke, 2005, p. 375). This study aimed to examine factors that may be barriers for physicians to effectively communicate risk to patients.
Using a theoretical framework based on Gagne’s learning theory, the mastery of higher-level skills requires prior mastery of lower-level skills (Gagne & White, 1978), understanding risk, a prerequisite skill of communicating risk, was explored. Additionally, in accordance with Gagne’s theory that posits learning outcomes fall in multiple domains including attitudes, the relationship between attitudes toward statistics and risk literacy was evaluated. Finally, Gagne explains instruction as the first step in a three-part learning process, hence the impact of prior statistics instruction on risk literacy was assessed.

This study confirmed that many medical students lack adequate risk literacy, a foundational skill needed to interpret and effectively communicate risk (Caverly et al., 2015; Trickey et al., 2014). As expected, attitudes toward statistics, both to the field and to the course, were positively correlated to risk literacy. Having successfully completed a college level statistics course did not improve participant performance on the risk literacy assessment; however, the small number of participants who did not fulfill this prerequisite may have led to the inconclusive findings. These results identify a competency deficit that needs to be addressed in medical school curricula.

Physicians are expected to interpret and understand complex risk and benefit information related to diagnoses and treatments; but the current training for physicians does not appear to adequately prepare them to master this objective. Furthermore, physicians are expected to effectively communicate risk to patients in a way that even those with low health literacy and numeracy can understand to promote patient engagement and shared decision making. Several underlying issues serve as barriers to this process. First, this study demonstrated that many medical students do not enter their training with the prerequisite skill of adequate risk literacy, despite having had a college level statistics course. Second, the literature did not reveal ample
coursework in medical school curricula that addresses this deficit; rather students are instructed in biostatistics and expected to already have basic statistic and probability knowledge. Third, even when physicians understand risk, communicating that risk to patients is a separate skill that merits training and placement in medical school curricula.

Governmental agencies and organizations have called for curricular change to integrate health literacy coursework in health professions training (Koh et al., 2012; Nielsen-Bohlman et al., 2004, p. 161; U.S. Department of Health and Human Services, 2010b, p. 28), but this mandate has not been implemented. Moreover, the LCME of the AAMC requires medical school curriculum to include health literacy training (Ross, Lukela, Agbakwuru, & Lypson, 2013), but the actual time dedicated to this competency is lacking across U.S. institutions. The content of that limited health literacy instruction delivered in medical schools does not specifically focus on health numeracy and communicating risk. Even physicians who complete their training with a strong aptitude for interpreting risk may never learn how to effectively communicate this information in a way that patients can understand to make informed decisions. Physicians may avoid shared decision making due to their own lack of risk literacy and/or the ability to explain risk to their patients.

Those responsible for medical school curriculum are challenged by an already crowded curriculum to find a place for what some consider a “soft skill” unlike basic science topics that are considered essential in the coursework. Certainly, one could argue that medical student risk literacy is just as necessary as the competencies of anatomy and physiology, as understanding risk guides treatment decisions. At a minimum, medical school curriculum should incorporate instruction related to basic statistics and probability needed to interpret risk in the field of medicine. Additionally, to better meet the needs of patients and satisfy national calls to action,
medical school curriculum should include instruction on health literacy and numeracy best practices for communication with a focus on communicating risk.

**Limitations**

High response rates at both Universities A and B, 98.2% and 98.7%, lead to the assumption that the medical students who voluntarily completed the survey are a representative sample of all first year medical students at their given institution; yet this study has limited generalizability of results due to a convenience sample representing only two institutions. Another limitation is that participants may have provided random answers without working the math problems since it is voluntary and does not count as a grade. To encourage effort to work out the four math problems on the Berlin Numeracy Test and to honestly answer the attitudes questions, the researcher stressed the importance of the study to improve medical school curriculum, which may result in improved patient health. Also, the results were consistent with previous research on statistical literacy and attitudes of health professionals and health professions students.

The most notable limitation in data analysis was the presumption that most students at University B would not have successfully completed a college level statistics course since it was not a prematriculation requirement; however, 89.0% of that group reported they did. This resulted in disproportionate groups for the predictor variable for this measure. The data showed a negative correlation between successful completion of a college level statistics course and score on the Berlin Numeracy Test which was not expected, but this correlation was not statistically significant. The results showed that even students who completed this prerequisite instruction did not demonstrate adequate risk literacy.

Although the data analysis in this study resulted in statistical significance in three of the
four regression models allowing for rejection of the corresponding null hypotheses, another limitation was the scoring of the Berlin Numeracy Test. A multiple regression was appropriately selected based on the types of variables being explored and the nature of the research question. Score on the Berlin Numeracy Test was used as the criterion variable and fits the definition of a ratio variable, an interval variable with an absolute zero; however, the only possible scores were 0, 1, 2, 3, or 4. A numeracy assessment with more questions and a broader possible score range may be a better tool to show a linear relationship with predictor variables but would also require more time for participants to complete.

**Recommendations for Future Research**

The results of this study demonstrated a deficit in first year medical student risk literacy as well as a correlation between students’ attitudes toward both the field and course of statistics to their performance on the risk literacy assessment. This research was only an initial step to determine medical school curricular needs and guide decisions on integrating competencies of risk literacy and statistical numeracy as well as effective methods for communicating risk to patients. Additional research to consider includes:

1. Examining risk literacy and attitudes toward statistics with third or fourth year medical students or residents to see if the skill is learned along the way despite not being represented in the formal curriculum.

2. Logistic analysis of the current data set using a cut score where a 3 or 4 is coded as adequate and 0-2 is coded as inadequate risk literacy.

3. Measuring risk literacy using an instrument directly related to the types of risk interpretation required by physicians rather than a general risk literacy measure.

5. Developing and testing risk literacy training for first year medical students.

6. Developing and testing risk communication training using health literacy and numeracy best practices for medical students after the first year but prior to patient encounters.

Although this study lacked balanced groups of medical students who had and had not successfully completed a college level statistics course, there would be no valid reason to repeat exploration of that relationship with risk literacy since this study established a majority of students who did complete the course still lacked adequate risk literacy. A college statistics course as a prematriculation requirement did not result in a medical students having strong statistical numeracy in terms of risk. Given the deficit in practicing physicians’ statistical literacy established over three decades ago, research efforts should primarily be focused on developing and testing medical school training and coursework as stated in the last two recommendations.
REFERENCES


Case Western Reserve University. (2016). *School of Medicine- University program and CCLCM program pre-requisite requirements*. Retrieved from https://case.edu/medicine/admissions/application-process/requirements/academic-pre-requisites/


The Family Medicine Milestone Project. (2014). *Journal of Graduate Medical Education, 6*(1s1), 74-86. doi: http://dx.doi.org/10.4300/JGME-06-01s1-05


Garcia-Retamero, R., Wicki, B., Cokely, E. T., & Hanson, B. (2014). Factors predicting surgeons’ preferred and actual roles in interactions with their patients. Health Psychology, 33(8), 920-928.


Han, P. K. (2013). Conceptual, methodological, and ethical problems in communicating uncertainty in clinical evidence. *Medical Care Research and Review, 70*(1 suppl), 14S-36S.


Karsenty, C., Landau, M., & Ferguson, R. (2013). Assessment of medical resident’s attention to the health literacy level of newly admitted patients. *Journal of Community Hospital Internal Medicine Perspectives, 3*(3-4).


Kumar, D., Sanders, L., Perrin, E. M., Lokker, N., Patterson, B., Gunn, V., ... Rothman, R. L. (2010). Parental understanding of infant health information: Health literacy, numeracy,


Oregon Health & Science University. (2016). *School of Medicine MD program- Requirements for admission*. Retrieved from http://www.ohsu.edu/xd/education/schools/school-of-medicine/academic-programs/md-program/admissions/requirements.cfm


Yale University. (2016). School of Medicine- Pre-medical requirements. Retrieved from https://medicine.yale.edu/education/admissions/apply/premed.aspx


APPENDICES

APPENDIX A: Berlin Numeracy Test

1. Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)? ________

2. Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in the choir 100 are men. Out of the 500 inhabitants that are not in the choir 300 are men. What is the probability that a randomly drawn man is a member of the choir? ___________ %

3. Imagine we are throwing a loaded die (6 sides). The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws, how many times would the die show the number 6? ___________

4. In a forest 20% of mushrooms are red, 50% brown and 30% white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with probability of 5%. What is the probability that a poisonous mushroom in the forest is red? _____________%
APPENDIX B: Permission to Use Berlin Numeracy Test

http://www.riskliteracy.org/researchers/

Use the Berlin Numeracy Test

The Berlin Numeracy Tests are fast user-friendly psychometric assessment technologies (e.g., measurement instruments), validated for use with educated samples from diverse countries and cultures (e.g., college students, computer-literate adults, physicians). The simplest Berlin Numeracy Test is a traditional 4-question paper and pencil test that takes < 4 minutes to complete. The computer adaptive version of the Berlin Numeracy Test takes about 2 minutes to complete because it only requires 2-3 questions that are selected based on participant performance. If a test-taker answers the first question right or wrong then a harder or easier question is automatically presented. We have also validated multiple choice formats, parallel forms, extensive full-scale and sub-scale tests (e.g., numeracy for certainty v. uncertainty), as well as very fast single-item tests for use with general community or highly-educated samples (i.e., median-split). All test formats are designed to address psychometric limitations of other numeracy and skilled decision making tests (e.g., negative skew, construct validity). A growing body of research indicates that the Berlin Numeracy Test tends to be the most efficient standalone assessment of numeracy, risk literacy, and general decision making skill currently available, more than doubling the predictive power of much longer numeracy and cognitive ability tests (e.g., intelligence, cognitive reflection, working memory; Cokely et al., 2012). We've also validated simple systems to combine our tests with other instruments for more extensive analyses, which can be valuable when working sub-samples like less-numerate patient groups. For help selecting the best test format for your needs please use our test recommendation tool.
APPENDIX C: Attitudes Toward Statistics Scale

ATTITUDES TOWARD STATISTICS
APPENDIX D: Permission to Use Attitudes Toward Statistics scale

Memorandum

To: Researchers Requesting the Attitudes Toward Statistics Scale
From: Steven Wise
RE: Instructions for using the ATS

You have requested a copy of the Attitudes Toward Statistics scale to use in your research. I have attached a copy of the scale that should be suitable for photocopying.

The 29-item ATS has two subscales. The Attitudes Toward Field subscale consists of the following 20 items, with reverse keyed items indicated by an “(R)”:

1, 3, 5, 6(R), 9, 10(R), 11, 13, 14(R), 16(R), 17, 19, 20(R), 21, 22, 23, 24, 26, 28(R), 29

The Attitudes Toward Course subscale consists of the following 9 items:

2(R), 4(R), 7(R), 8, 12(R), 15(R), 18(R), 25(R), 27(R)

To score the ATS, simply sum the appropriate item scores for the subscales and/or total scale.

The original reference for the ATS is:


Another useful (and more recent) reference regarding the scale is:


In exchange for permission to use my scale, I'd appreciate your sending me a copy of any manuscripts that result from your use of the ATS, as I am always interested in seeing studies that use the scale. I thank you in advance for your cooperation, and I wish you success with your research.

Harrisonburg, Virginia 22807
(540)388-8708
(540)388-7878 Fax
APPENDIX E: Email Request to Include University A Students in Research

Dear Dr. ,

Hello. I am a doctoral candidate at Liberty University and also the Director of Programs and Training at the Center for Health Literacy at the . My director, Dr. , suggested I contact you regarding my research plans. For my dissertation, I am wanting to research health numeracy and focus specifically on first year medical students’ risk literacy and attitudes toward statistics to identify a possible gap in the curriculum that may be a barrier to effective patient-provider communication. Please review the attached abstract that outlines my general plan. I am requesting your permission and cooperation to survey the first year college of medicine students at . Could we please set up a time to collaborate in person or via a conference call to discuss additional details and logistics? I can be reached at . Thank you for considering my request. If you choose to grant permission, please respond to me by email.

Sincerely,

Tina D. Moore
Dear Dr. [Name],

Hello. I am a doctoral candidate at Liberty University and also the Director of Programs and Training at the [Name of Institution] Center for Health Literacy at the University of Arkansas for Medical Sciences (UAMS). Your colleague, Dr. [Name], suggested I contact you regarding my research plans. For my dissertation, I am wanting to research health numeracy and focus specifically on first year medical students’ risk literacy and attitudes toward statistics to identify a possible gap in the curriculum that may be a barrier to effective patient-provider communication. As you know, Dr. [Name] is highly involved in health literacy research. He believes Oregon Health and Science University would be an appropriate setting for my dissertation research. Please review the attached abstract that outlines my general plan. I am requesting your permission and cooperation to survey the first year college of medicine students at [Location]. Could we please set up a time to collaborate via a conference call to discuss additional details and logistics? I can be reached at [Contact Information].

Thank you for considering my request. If you choose to grant permission, please respond to me by email.

Sincerely,

Tina D. Moore