12-2015

GME Graduate Retention Rates: A Single Institution Study

Tracy J. Frieswyk
Western Michigan University, tfrieswyk@yahoo.com

GME GRADUATE RETENTION RATES: A SINGLE INSTITUTION STUDY

by

Tracy J. Frieswyk

A dissertation submitted to the Graduate College
in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy
Educational Leadership, Research and Technology
Western Michigan University
December 2015

Doctoral Committee:

Jessaca Spybrook, Ph.D., Chair
Chris Coryn, Ph.D.
Alan T. Davis, Ph.D.

GME GRADUATE RETENTION RATES: A SINGLE INSTITUTION STUDY

Tracy J. Frieswyk, Ph.D.

Western Michigan University, 2015

Graduate medical education (GME) refers to the advanced instruction provided to clinicians who have previously received their MD (doctor of medicine) or DO (doctor of osteopathic medicine) degrees. GME education takes place in the clinical setting (e.g., hospitals, clinics), delivering the training necessary for physicians to become licensed to practice medicine, as well as to become board-certified in their specialty. GME training programs throughout the US are absolutely essential, as they are the primary source in this country for the physician workforce.

One of the major decisions in a physician's life as they approach the end of GME training is where to practice medicine. Federal and state governments, along with other sources devote substantial resources to the training and development of medical doctors. In a time of increased concerns over an impending shortage of physicians and the costs associated with GME training, there are key incentives to identifying those GME graduates who are most likely to practice medicine in the state in which they trained. Thus, GME training programs are interested in learning more about the factors

that influence in-state practice location decisions, as well as how to identify graduates that are likely to practice in-state.

The focus of this dissertation was to utilize logistic regression with cross-validation to examine in-state retention using individual level demographic and educational predictors in order to create a pilot-scoring tool to identify graduates from a Michigan-based GME training program who are likely to practice medicine in Michigan post-training. Results showed that a connection to the state of Michigan (e.g., being born in Michigan, graduating from a university or medical school in Michigan and completing GME training in Michigan), as well as graduating from a primary care program and being married, were predictive of in-state retention. A score associated with each variable was determined and a pilot-scoring tool was created to identify GME graduates likely to practice in Michigan post-training. A tool like this could be used in targeted recruitment efforts towards graduates likely to practice in Michigan after training. Further studies to determine the reliability, validity and applicability of this scoring tool are necessary.

ACKNOWLEDGMENTS

First, I would like to thank my boss and committee member, Alan T. Davis, PhD, as well as the other members of my committee, Jessaca K. Spybrook, PhD and Chris Coryn, PhD, for assisting, mentoring and guiding me through this process. If it were not for Dr. Davis I am not sure I would have made it through the EMR program/dissertation. His ability to listen to my frustrations and complaints about feeling like I would never understand anything all the while patiently mentoring me through school, dissertation and work has been an invaluable experience to my development as a student, employee and human being!

This goes for Dr. Spybrook as well. She helped keep me sane throughout the dissertation process, while I was probably driving her insane. Her ability to turn the reviews of my chapters around on a dime helped me get through my dissertation much more quickly than I could have ever hoped for. Also, she is an outstanding educator and mentor and overall amazing person. Dr. Coryn also has an outstanding ability to educate and mentor, I learned so much from the classes he taught.

To my friend Sarah, you rock! Thank you for all of the writing sessions and meals that we shared together. Two friends bonded for life by going through dissertation at the same time. You made the process much more fun than going through it alone! A special thank you to Jaclyn Goodfellow, Director of GME at GRMEP, and all the

Acknowledgements - Continued

program coordinators and other GRMEP staff that assisted with my data collection efforts and answered my many questions! Thanks for the support of my daughter, Ashley, and husband, Henry, as well as the rest of my family. Henry, thank you for all of your support, encouragement, understanding and patience while I missed dinners, family functions and other important events to write papers, study and work on my dissertation – I couldn't have done it without you!

Tracy J. Frieswyk

TABLE OF CONTENTS

Table of Contents - Continued

CHAPTER

Table of Contents - Continued

CHAPTER

Table of Contents - Continued

CHAPTER

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Statement of the Problem

One of the major decisions in a physician's life as they approach the end of graduate medical education (GME) training is what to do next. While a small minority of doctors go into industry or pursue full time research, the vast majority of physicians will practice medicine. Then the question becomes, where to practice? The federal government and GME sponsoring institutions devote substantial resources to the training and development of medical doctors. In a time of increased concerns over an impending shortage of physicians and the costs associated with residency training, there are key incentives to identifying those residents who are most likely to practice medicine in the state in which they trained (i.e., in-state retention). Thus, GME training programs are interested in learning more about the factors that influence in-state practice location decisions, as well as how to identify graduates that are likely to practice in-state.

This dissertation utilized techniques from the field of research, logistic regression with cross-validation, to examine the demographic and educational characteristics that may influence the decision to practice medicine in Michigan. After the predictors from the regression were identified, a pilot-scoring tool to aid in the evaluation of whether graduates were likely to practice within the state of their GME training was developed. As there are currently no tools like this available, an instrument like this is a valuable

contribution to the field of evaluation. It also contributes to the field of GME as it would allow sponsoring institutions to identify whether upcoming graduates are likely to practice medicine in the same state in which they trained. This tool could be used in targeted recruitment efforts of these identified individuals.

What is GME?

The educational journey to becoming a physician spans many years. It begins with an undergraduate degree followed by four years of medical school or undergraduate medical education (UGME) and then another three to 11 years of GME.[1] This type of education involves caring for patients, attending educational conferences and didactic sessions, learning new skills and techniques, participating in scholarly activities (e.g., research) and working with other health professions learners. GME delivers the training necessary for physicians to become licensed to practice medicine, as well as to become board-certified in their specialty. GME training programs throughout the US are absolutely essential, as they are the primary source in this country for the physician workforce.[2]

The length of GME training is based upon the specialty the physician is pursuing.[1] For example, an internal medicine residency program lasts three years, while a general surgery program requires a five year commitment. Sub-specialty training (e.g., vascular surgery, pediatric hematology/oncology), in the form of fellowship programs, can add another one to three years of additional GME training. As a fellow, the physician's role is basically that of an attending physician (i.e., a board certified physician who is in

practice and seeing patients) in the area of whatever specialty in which they completed residency, while learning a more specific sub-set of knowledge within that specialty.

GME Training Sites

The bulk of GME training takes place in teaching hospitals and ambulatory offices under the supervision of practicing faculty physicians. Teaching hospitals make up approximately 6% of all of the hospitals (n=5,686) within the US.[2-4] Michigan alone has 44 GME sponsoring institutions (52 teaching hospitals), ranking 6th in the nation, supporting just under 5,000 medical residents per year.[5,6,7]

Teaching hospitals play a pivotal role in supporting physician training by providing clinical experiences to learners, as well as opportunities for scholarly activity in the form of research, patient safety and quality improvement projects.[1,8] In addition to providing a large amount of complex and acute patient care, one fifth or more of all of the care which takes place in US hospitals occurs at teaching hospitals.[1,2]

These training sites provide many advanced services with the latest technology that non-teaching hospitals cannot provide.[2,4] This includes burn units, pediatric and neonatal intensive care units, transplant services and cardiac surgery.[2] GME training sites not only provide many benefits in the short-term in the form of offering more extensive and advanced services, they are also the source of America's physicians for the future.

However, providing these resources does not come without a price. Residents and fellows receive salaries and benefits during their training period. Factoring in the indirect costs to the teaching hospitals, which are associated with residency training

(e.g., more tests ordered, longer patient length of stay), the average cost of GME training has been reported to average between $152,000 to $239,000 per resident or fellow per year.[1,9]

GME Funding

GME is funded by multiple sources, including government sources, as well as private sources. The two main government sources include federal funding in the form of Medicare and Medicaid. [1,4,10,11] Other sources include Veterans Affairs, Health Resources and Services Administration and the Department of Defense.[10,11] State governments are also able to provide additional sources of funding to GME. Private funding can come from the teaching hospitals themselves, philanthropy, private insurers, universities and faculty physician practice groups.[1,11,12]

GME Funding from Medicare

The main source of funding comes from Medicare payments from the federal government, which account for over $9 billion of the funds provided for GME per year.[4,10,11] There are two types of Medicare payments made to teaching hospitals, direct GME (DGME) and indirect GME (IGME).[1,11] DGME payments cover resident salaries and teaching faculty time. Since teaching hospitals are paid at the same discharge rate as non-teaching hospitals, and these rates are usually less than the cost of the service provided, congress proposed IGME payments to cover the increased expenses related to the operational costs of the GME programs (e.g., more tests ordered, longer patient stays, advanced services offered, complex patient population).[1]

Pressures on GME funding have been mounting over the last 20 years. In 1997, the Balanced Budget Act was introduced and passed by Congress.[1,8,11] It included a provision to reduce the amount of federal support for GME funding by capping the number of residency slots that can be supported by Medicare. In more recent years there have been several bills introduced in the US Congress to reduce GME payments. For example, a proposal of a 50% reduction in GME by the Joint Selection committee on Deficit Reduction was added to the Budget Control Act of 2011, however, this language was later omitted.[11]

GME Funding from Medicaid

Medicaid is the second largest source of GME funding, providing over $3.78 billion per year. [10,11] State governments can use Medicaid dollars to support GME programs, although there are no federal regulations on how a state chooses to disburse funding or if they choose to disburse funds to GME at all.[6,11] Most states have historically provided GME funding through Medicaid. The range of state Medicaid support is variable. In 2012, Alaska spent $375,000 and New York spent $1.8 billion, while eight states used no Medicaid funds to support GME.[6] This lack of regulation leads to wide variability in how Medicaid funds are used.

In more recent years, some states have reduced and/or limited the number of Medicaid dollars allotted for GME, while others have eliminated this type of financial support all together. For example, Michigan reduced Medicaid funding to GME by 3.6% from 2009 to 2012.[6,13] Michigan spent $163 million in Medicaid dollars on GME funding during 2012, averaging 3.1 million dollars per teaching hospital.[6] More recently, a $57

million cut to Medicaid funding for GME was proposed for the 2016 budget.[14] However, the current funding structure was maintained for the 2015-16 budget year.[15]

Implications of Cutbacks to GME Funding

Major cuts to GME funding may force some teaching institutions to reduce or cut GME programs and/or eliminate services that are unavailable elsewhere in the community.[1,4,16] These types of reductions, in an era when Americans are living longer and are requiring more complex care, may adversely affect the quality and availability of medical care in the US. The age group with the greatest healthcare needs are those >65, and the Census Bureau projects a 36% increase in this group by 2020 due to the aging baby boomers.[4,16,17]

To further add to the complexity of the situation, the Affordable Care Act (ACA) has introduced apprehension over reduced profit margins for hospitals.[18] In addition, demand for healthcare is also projected to increase with the implementation of the ACA.[4] Those who did not have health insurance before will be able to take advantage of services that were not readily available to them due to lack of coverage. Finally, the ACA has fueled concerns over the projected physician shortage, estimated to be 91,500 nationwide by 2020, with a shortfall of 6,000 projected for the state of Michigan.[2,17,19]

The strain on available health related services provided by physicians is further complicated by the federal cap, which limits the number of GME training positions funded by Medicare. This barrier to GME training has created a bottleneck related to the number of physicians entering the US workforce at any given time. An aging physician workforce, of which 1/3 are projected to retire by 2020, will further widen the

gap between healthcare needs and the number of physicians available.[4,16,17] In an effort to close the gap, medical schools have been popping up all over the US, including three in Michigan within the past three years.[8] However, the increase in medical school graduates will do nothing to offset the burden if there is no increase in government funding to support additional GME training slots, as well as maintaining the current funding structure of GME.

GME Graduate Retention and Return on Investment

There is a need for more transparency and accountability from the GME sponsoring institutions that receive the funds for their training programs.[11] The issue with the current model is that GME funds received through Medicare, Medicaid and other sources go directly into the revenue stream of the teaching hospitals.[11] This practice makes it difficult to get an accurate picture of how the funds for GME are actually being spent by these training sites. This practice has resulted in a lack of accountability to GME funding sources on the part of GME institutions. If the current funding structures are to be maintained, GME sponsoring institutions need to find ways to demonstrate to funding sources the value of their investment (i.e., return on investment (ROI)).

Several recent reports have focused on how the GME system is controlled and funded, and have also examined alternative funding sources and ways to demonstrate ROI for investment in GME.[10,12,20,21] Many of these reports suggest that tracking location of practice post-training could be used to demonstrate ROI. Location of practice after graduation could be in the form of rates of graduates who practice in health professions

shortage areas (HPSA) and/or in-state retention rates (i.e., practicing in the same state as GME training), both of which could demonstrate ROI to GME funders. In fact, legislation in Massachusetts directed a special commission on graduate medical education to look at the effect of GME on fulfilling physician needs in specific specialties, as well as to determine different ways to fund GME.[20] The final report recommended that performance benchmarks tied to GME funding could be in the form of in-state retention and training physicians in specialties that are linked with physician shortage projections, just to name a few.[22]

In a study to investigate the potential expansion of GME in Northwest Indiana, Tripp Umbach reported that physicians that practice in the state of GME training generate approximately $1.5 million per year in economic benefits.[21] If GME sponsoring institutions could demonstrate a high in-state retention, this could further justify funding for GME, especially in the form of Medicaid dollars, which come from the state. For example, if Michigan provides Medicaid funds towards GME and less than half of the GME trainees end up practicing in the state, legislators may not recognize this as a good ROI.

Since Michigan is one of 22 states that have linked the goal of increasing the physician workforce to Medicaid GME payments[6], the concept of in-state retention could be used as a marker for justification of funds. The ROI in this case would be closing the physician shortage gap leading to improved access to health care, as well as the economic benefit to the state and community where the physician is employed. In this scenario, there would be a need for Michigan-based GME programs to explore ways

to identify residents/fellows early in training who are likely to end up practicing medicine in Michigan. Targeted recruitment efforts aimed at these individuals could potentially increase in-state retention rates.

Study Purpose

In an era of limited financial resources, there is a growing need for GME sponsoring institutions to demonstrate to funding sources the value of their investment. One strategy for dong this is to increase their in-state retention rates. This would serve a dual purpose of justifying the expense of the physician education, as well as address concerns over physician shortage issues in the state. The purpose of this study is to examine individual level characteristics of graduates from a Michigan-based GME sponsoring institution to develop a tool to identify graduates that are likely to practice medicine in the state of Michigan.

Study Objectives

More specifically the first objective of the study is:

1. To use logistic regression with cross-validation to examine the individual characteristics related to whether or not graduates who trained in a Michigan-based GME sponsoring institution practice medicine in Michigan.

Further, using the identified predictors from the previous analysis to create a mechanism to help pinpoint graduates that are likely to practice in Michigan post-training would be useful to Michigan-based GME sponsoring institutions and/or physician recruiters in Michigan. A tool like this could also prove useful to GME

sponsoring institutions in other states. The second objective of the study will address this concept.

2. Create a scoring tool based on the logistic regression with cross-validation to categorize GME program graduates into groups: likely to practice in Michigan or not likely to practice in Michigan.

Contribution to Evaluation, Measurement and Research

Empirical evidence from a single Michigan-based GME sponsoring institution was used to examine variables related to whether or not a graduate practiced medicine in Michigan. Five-fold cross-validation and bootstrap techniques were combined with logistic regression, techniques from the field of research, to build and test the model prior to creation of the pilot-scoring tool. Variables that were identified as predictors of practice location were weighted and used to create a pilot-scoring tool that could be used to evaluate whether or not GME graduates are likely to practice medicine in Michigan.

This data driven approach to developing a pilot-scoring tool is a contribution to evaluation. The tool has the potential to advance the evaluative capacity of Michigan-based GME institutions to assess the likelihood a GME graduate would practice in Michigan. This tool could be used to produce a listing of identified graduates likely to practice in Michigan for hospitals and other physician recruiters in Michigan to use for targeted recruitment of these individuals.

Organization of Dissertation

There are five chapters included in this dissertation. The material in chapter one introduced the reader to the concept of GME and its importance. How GME is funded and the issues surrounding funding are also covered. The concept of in-state retention is described as well. Lastly, the statement of the problem, study purpose and objectives are described.

Data available in the literature on graduate retention within the GME training state are discussed in chapter two. The methodologies of this study are detailed in chapter three, including the study sample, data collection, assumptions of logistic regression and other analytical methods.   The results of the study are reported in chapter four. Chapter five includes a discussion of the study findings, implications of the study, recommendations for future research and conclusions.

CHAPTER II

REVIEW OF LITERATURE

This chapter includes a review of the available literature on GME retention. Also described are limitations of the methodologies used to obtain, analyze and report the data relating to retention. Studies that have been used to design scoring systems are also discussed. Lastly, a different methodological approach to examining in-state retention is described.

Physician Retention

Physician retention is an important part of the discussion related to funding for GME training. It has been stated that GME training programs that produce physicians that return or stay to practice within the state of GME training is economically beneficial and increases the access of the public to healthcare.[2,16,21] Retention of physicians in the state in which they completed GME training, as well as in physician shortage/rural areas, have been proposed as performance measures to demonstrate ROI.[22]

Predictors of retention of physicians in physician shortage/rural areas and in-state retention have been studied. First, the literature related to predictors of retention in physician shortage/rural areas is reviewed. Next, studies focusing on examining factors related to retaining physicians within the state of GME training are discussed, as well as additional published reports that include data on in-state retention rates of physicians who practice medicine in the state in which they received their GME training.

Physician Recruitment/Retention into Physician Shortage/Rural Areas

With the impending physician shortage issues, access to health care, especially in rural settings, is an important component of the retention conversation. Many studies have focused on examining factors related to retaining physicians to practice in physician shortage/rural areas.[23-29] Due to the lack of resources and health disparities (e.g., older, sicker, poorer) associated with rural settings, recruitment and retention of primary care physicians in rural areas has been a topic of investigation over the past 100 years.[23]

More recently, studies to investigate factors related to choosing primary care, as well as practicing in rural and underserved communities, have been conducted.[23-29] These reports include the use of historical data, as well as data from surveys, to examine the factors that influence the choice to practice medicine in rural areas.[23-26] Other studies have included various methods to obtain data (e.g., semi-structured interviews, document review, observations) in order to investigate spousal and community roles, as well as other factors related to recruitment and retention into a rural practice setting.[27-29] These studies are discussed in the following paragraphs.

Rabinowitz et al. conducted a retrospective study of Jefferson Medical College (JMC) graduates from 1978 to 1993, including a subset of graduates that participated in a Physician Shortage Area Program.[23] One purpose of the study was to determine predictors of rural primary care physician retention. This study utilized data previously collected for a longitudinal study tracking JMC graduates, as well as from the American Medical Association (AMA) Physician Masterfile.

There were 19 study variables investigated by the authors from the following areas: demographic, pre-medical, career plans in medical school, medical school programs/curricula and economic issues. Univariate analyses for each of the 19 variables were performed to examine their relation to rural practice. Significant findings from the univariate analyses were entered into a multivariate model to assess their predictive ability related to practicing in a rural setting. Significant predictors from the multivariate regression analysis included plans to practice in family medicine during the freshman year of medical school, participation in a physician shortage program, National Health Service Corps (NHSC) scholarship recipient, rural preceptorship experience and male gender. The authors also found that individuals that had a rural background, combined with a plan to practice in family medicine (developed during the freshman year of medical school), were more than twice as likely to practice in a rural setting than individuals with only one of these variables.

Another study, conducted by the Robert Graham Center and funded by the Josiah Macy Jr. Foundation, examined factors that influence medical student and resident choices about medical specialties (primary care vs non-primary care) and location of practice (rural and underserved populations).[26] Specifically, the roles that student debt, scholarship and loan programs, type of school, curriculum, institutional culture, and potential income have on medical students' specialty choices, as well as decisions to care for underserved populations.

Data were compiled from survey results from medical students completed at graduation and historical data related to receipt of Title VII funds during training, the

AMA Physician Masterfile and participation in the NHSC, as well as from data related to current practice specialty and location, and service in Rural and Federally Qualified Health Centers. Data from a total of 322,131 individuals were available for analysis.

Stepwise logistic regression analyses were used to examine a variety of independent variables and their relation to the dependent variables of practicing in a rural setting and practicing in a physician shortage or underserved area. Given the large sample size, it was clear that most, if not all, of the independent variables would be statistically significant, so the investigators put the added restriction that the most important variables would have an odds ratio (*OR*) >2 or <0.5. For example, in their analysis of likelihood to practice in a rural area, 20 of the 21 independent variables were statistically significant ($p<0.05$), while only four of the variables were considered to be important predictors.

Chan et al. also examined variables related to the decision to practice rural medicine. A survey was conducted of rural family physicians who graduated from Canadian medical schools between 1991-2000.[24] A total of 382/651 physicians (58.7%) completed the questionnaire. Comparisons were made between physicians with an urban upbringing and a rural upbringing with regard to rating important factors that influenced the decision to practice rural medicine, using the chi-square test. Statistically significant differences were seen with regard to educational training (rural exposure during medical school/residency), rural exposure growing up and other factors including proximity to family, spouse/partner influence, desire to practice where need the greatest.

A survey to assess aspects of the environment that could be used to predict retention to use in the recruitment of physicians to the Hawaiian Islands was conducted by Pellegrin.[25] Hawaii suffers from physician shortage issues, therefore the author felt it was important to study factors related to recruitment and retention of physicians to the island. There were 127 participants in the survey. Predictors of retention included access to good K-12 school systems, financial sustainability, community support and professional opportunities.

Other studies have taken a more qualitative approach to the issue of recruitment and retention of physicians into rural medicine.[27-29] One study by Hancock et al. used data collected from 22 semi-structured interviews to create a model consisting of "four main pathways to successful and fulfilling rural practice: familiarity, community, sense of place, and self-actualization" (pg. 1372).[27] These 22 interviews were conducted with mostly white males, with an average age of 55 years. Mayo et al. conducted 13 interviews of spouses of rural physicians to "gain a better understanding of spousal concerns and experience regarding rural living" (pg. 272).[28] Cameron et al. used a case study design to study four rural communities to "explore the professional, personal and community domains of physician retention" (pg. 47).[29]

In-State Physician Retention Data

Over the years, in-state retention has been examined in many different ways. These include studies that used logistic regression to examine variables related to in-state retention, as well as studies that report in-state retention of specific program (e.g., family medicine) graduates or summary data related to percentages of graduates who

practice in the state in-which they underwent GME training.[9,30-35] These studies are discussed in more detail in the following paragraphs.

A study in 1986 by Burfield et al. using data from the AMA Physician Masterfile examined the relationship between medical training and practice location.[30] The authors included personal (e.g., age, time since graduation) and professional (e.g., practice specialty) characteristics of physicians, as well as medical school (e.g., US vs non-US, reputation) and state (e.g., population, per capita income) characteristics for active physicians in 1982 in their study. Summary data were reported. When specifically looking at in-state retention of GME graduates, the results showed that a higher percentage of general/family practitioners, as well as female physicians practiced in the state in which they completed GME training. Overall, 51.1% of physicians were practicing in the state in which they received GME training and 29.7% completed both undergraduate medical education (UGME) and GME in the state.

In 1995, Seifer et al. published a study that also used the AMA Physician Masterfile to assess the relationship between GME and practice location, specifically location in the state in which GME training occurred.[31] The investigators took a random sample from the 1993 edition of the AMA Physician Masterfile and coupled it with the American Osteopathic Association (AOA) Physician Masterfile for the analysis. Summary data were reported. A logistic regression was also performed to examine the relationship between physician/state characteristics and practice location in the state of GME training. The authors reported that in 1993, 51% of physicians were practicing in the state in which they trained. Data were also reported for each state. Michigan

showed an in-state retention rate of 48%. The logistic regression analysis revealed that gender (female), UGME in the same state as GME training, graduating from a US medical school (inverse relationship), specialty (generalist), professional activity (teaching, research, administration), board certification (inverse relationship), a federal employee (inverse relationship), number of resident physicians in the state (inverse relationship), number of non-resident physicians in the state and percentage of the population being urban were predictive of in-state retention, while age, professional degree (MD vs DO) and state income level were not.

Another study using data from the AMA Physician Masterfile analyzed the effects of birth location, medical education and completion of GME training on practice location for family medicine residents.[32] The study sample included physicians that completed GME between 1997 and 2003 that were born in Virginia, attended medical school in Virginia or completed GME training in Virginia. Seven different variable combinations were analyzed to assess the likelihood of practicing in Virginia. A total of 806 physicians met the inclusion criteria. The results for each variable on its own were reported. Only 6% of physicians who were born in Virginia but received UGME and GME training in another state were practicing in Virginia. There were 17% of physicians practicing in Virginia who had attended medical school there, but were not born in the state nor did they undergo GME in the state. Data for physicians who received GME training in the state, but were not born in Virginia and did not attend medical school in the state were reported to be 49%. Combinations of variables revealed that 74% of physicians practicing in Virginia were born in the state and completed GME in the state,

while 82% went to medical school in the state and completed GME training in the state. Finally, 81% of the physicians who had the combination of born in the state, attended medical school in the state and completed GME training in the state were identified as practicing medicine in Virginia.

Fagan et al. also utilized the AMA Physician Masterfile to assess the relationship between the location of practice and where the physician received their family medicine training.[33] The investigators examined practice location within 5, 25, 50, 75 and 100 miles of GME training, as well as the percentage of family medicine graduates practicing in the state in which they trained. A total of 64,972 physicians were included in the study. The results showed that, nationally, 57% of the family medicine graduates practiced in the same state as their GME training site and 55% within 100 miles of the training site. Individual rates were reported for each state as well.

Another source of in-state retention rates that uses data from the AMA Physician Masterfile comes from the Center for Workforce Studies, which publishes a biennial report on GME in the US, physician supply, and medical school enrollment.[5] Included in this report are data on rates of GME graduates who practiced within the state in which they received their GME training for active physicians (e.g., administration, direct patient care, medical research, medical teaching).

The main sources of data for the 2013 report are listed as the AMA Masterfile (December 2012 file), population estimates from the US Census Bureau, Association of American Medical Colleges (AAMC) Student Record System, AOA and National GME Census (conducted by the AAMC and AMA).

The workbook reports data for each state, which are then re-reported in various reports and briefs by individual health healthcare workforce committees. For example, the Office for Healthcare Workforce Analysis and Planning, based in South Carolina, published a brief in 2013 on retention of physicians who trained in the state.[34] The data in this brief were based on the 2011 State Physician Workforce Data Book published by the Center for Workforce Studies. It was reported that, of the active physicians in 2010, 2,389 attended medical school and residency training in South Carolina. Of those physicians, 1,829 (76.6%) were practicing medicine in South Carolina.

In-state retention data were also published for the Georgia Statewide Area Health Education Center (AHEC) Network in relation to increasing the number of primary care GME slots in Georgia. Data reported indicated that about 74% of physicians who graduated from a Georgia-based high school and GME training program practice in Georgia. Further, over 80% of physicians who graduated from a Georgia-based high school, medical school and GME program remained in the state to practice.

Another report by the Center for Workforce Studies focused on GME graduates in New York.[35] For the past 15 years during the Spring prior to graduation, an annual exit survey has been given to residents and fellows completing their training in the State of New York. The survey results report information that includes post-training plans, specialty needs and job prospects in the state of New York.

The survey is distributed to GME sponsoring institutions located in New York. It is assumed that the GME institution will then send the survey to graduating residents and fellows. The overall response rate for the most recently published year (2014) was 56%

(2,951/5,275). Across the past 15 years, the overall response rate has been 61%. Data collected in the survey span four categories: demographic/background characteristics, post-graduation plans, future employment plans, job search/job market experiences.

Results from the most recent survey showed that, of those that responded, just over half of the graduates were planning to enter the patient care setting upon graduation, and of those, 83% had confirmed practice plans. Of the graduates who had confirmed practice plans, just under half (45%) were planning on practicing in New York. A high percentage (80%) of the respondents who were staying in New York had strong ties to the state, attending both high school and medical school in New York. Respondents who were leaving New York to practice gave reasons such as: to be closer to family (27%), better jobs (14%) and salary (9%), no intention to stay after training (6%), climate/weather (2%) and taxes (1%).

Methodological Considerations of the Published Literature on Retention

AMA Physician Masterfile Studies to Examine Retention

While the data from the published literature provide useful information on retention, the majority of these studies are primarily based upon data from the same source, the AMA Physician Masterfile.[23,26,30-35] This data source has been maintained since 1906 and is the source of data for over 1.4 million physicians, residents and medical students in the US. The AMA has a division (Division of Surveys and Data Resources) that manages the data file.[36,37]

Data for the AMA Physician Masterfile are sourced from medical schools, GME programs, state licensing agencies, disciplinary action reports from medical and osteopathic boards, National Board of Medical Examiners reports, Educational Commission for Foreign Medical Graduates, American Board of Medical Specialties, and Federal Drug Enforcement Administration registration. A request to AMA members to update their profiles is made every three years. Other non-member profiles are updated using the other sources listed.

One concern with the data from the source is the accuracy at any given point in time. One study showed that the AMA Physician Masterfile data overestimated the number of physicians practicing in small rural communities.[38] The three-year lag time between updates, as well as the reliance of physicians who belong to the association to update their information may lead to inaccuracies in the data. For example, one graduate included in the data for this dissertation was still showing a Grand Rapids, Michigan location (left over from residency) using this resource, when it is known that this individual had not practiced in Michigan since graduation.

Regression to Examine Retention

A small number of studies used regression to assess predictors of practice location within physician shortage/rural areas, as well as within the state of GME training.[23,26,31] These studies provide useful insight as to factors that influence practice location. However, there are limitations to some of the methodology used to examine retention. There are also differences between the factors used to examine retention as

defined for this dissertation and those examined in other studies. These limitations and differences are discussed in the following paragraphs.

The study of Rabinowitz et al. used multivariate analysis to determine optimal predictive factors for rural supply and retention of physicians.[23] One concern is the method by which the variables were selected for inclusion in the model. The investigators set their significance level cut-of at 0.05 for the univariate analyses as being the criterion for inclusion in the multivariate analysis. However, this low of a significance level can lead to incorrectly rejecting variables that should be included in the regression model. Sun et al. showed that variable selection using this method can lead to incorrectly rejecting variables to include in the model, which may occur due to confounding with other variables.[39] Harrell has also expressed concerns with this methodology for variable selection.[40]

Seifer et al. utilized logistic regression to examine factors related to in-state retention of GME graduates, which was the same outcome variable used for the current study.[31] However, they included factors related to physician/professional characteristics and state-level characteristics in the analyses, many of which were different than the factors examined in this dissertation. More currently, the Robert Graham Center used regression to examine factors that influence medical student and resident choices related to practicing in rural and underserved areas, as well as practice specialty choices.[26] However, this more recent study differs from the current dissertation as in-state retention rates of GME graduates was not the focus.

Survey Studies to Examine Retention

Multiple factors related to error (e.g., coverage, sampling, non-response, measurement) are associated with survey studies.[41] Error related to coverage can be introduced if the entire target population was not included based on the method of survey dissemination. For example, under-coverage could be an issue if the entire target population does not have an opportunity to complete the survey. Conversely, over-coverage can be problematic if duplicates and/or those other than the targeted group have an opportunity to complete the survey. Non-response bias can be a limitation to the interpretation of survey results if response rates are low. Measurement error can also be introduced through the instrument itself through poorly worded questions and poor design/layout.

In the case of the exit survey of New York GME graduates, response bias may be an issue.[35] In this case, under-coverage is an issue since there is an assumption made that GME sites are sending the survey to their graduates. For the survey study conducted by Pellegrin, it is unclear whether under-coverage or over-coverage may be a concern.[25] For this study, key healthcare leaders and other physicians were asked to distribute a link to the survey via email to clinicians in their associated institutions. It would be unclear from this method of survey distribution whether or not everyone in the target population received that email or not. These are limitations of survey studies that rely on others for distribution, however, sometimes this is the only option of dissemination available to investigators.

Non-response bias may be an issue for many of these studies since their response rates were less than 100% and random sampling was not employed.[24,25,35] Pellegrin reported that the response rate is unknown for her study due to the method of distribution of the survey.[25] Chan et al. reported a response rate of 59%.[24] However, the reported rate may be problematic since the initial mailing included 784 physicians and 133 returned questionnaires were deemed to be ineligible for various reasons. It is unknown how many other surveys were sent to ineligible candidates (over-coverage), so the true response rate can never be known.

One last limitation of the New York GME survey is that the in-state retention data are only reported for those with confirmed practice plans at the time of survey dissemination.[35] There is no confirmation that those physicians ever ended up practicing in the state. Further, those without confirmed practice plans who completed the survey may have ended up practicing in New York. Another missing component of the in-state retention data are the percentage of non-responders who stayed in New York to practice medicine.

What's Missing from the Published Literature?

What seems to be lacking in the literature is the contribution of specific GME training sites to in-state retention rates. Further, a more current look at predictors, specifically demographic and educational characteristics of the physician, of in-state retention is needed. Lastly, a method to identify graduates likely to practice within the state of GME training, such as a scoring tool, would be beneficial to GME sponsoring institutions, as well as hospital and physician recruiters.

Scoring Systems Developed through Multivariate Regression

Many studies have been designed to create scoring systems to predict specific outcomes.[42-45] The ability to predict an outcome is useful in many ways. In medicine, scores developed from multivariate analyses to predict one year survival in ICU patients on a mechanical vent or risk of heart dehydration in children with diarrhea are designed to assist physicians with treatment decisions. In education, scores can be used to assist with making decisions about educational interventions designed to retain students. Methods used to develop these scores include the use of multivariate predictor models coupled with cross-validation techniques.

Hough et al. developed a system to predict mortality for adult ICU patients on mechanical ventilators for prolonged periods of time.[42] The model was developed using retrospective data captured on day 14 for patients receiving mechanical ventilation, from 40 US medical institutions. A stepwise logistic regression analysis was employed to develop the model. A development cohort (n=491) and validation cohort (n=245) were used to create and test the model. The investigators used the area under the curve (AUC) developed from a receive operator characteristic (ROC) analysis and Hosmer and Lemeshow's goodness-of-fit statistic for model evaluation. Prior to developing the scores for the model, continuous variables were turned into categorical variables and another logistic regression was performed on the development cohort to obtain the β-coefficients to use in score development. Scores were ranged from zero to four or greater. Kaplan-Meier analyses were performed to plot survival for each score category for both the development and validation cohorts to assess performance.

Zodpey et al. created a risk-scoring system to assess children with dehydration related to diarrhea.[43] A multiple regression analysis was conducted using data from 774 patients. A total of 17 hypothesized risk factors were tested in the model. Factors in the final model were weighted and each weight was rounded to the nearest whole number after multiplying it by 10. Scores for each of the patients in the sample were then determined. The investigators then calculated the sensitivity, specificity, and predictive accuracy. The best cut-off score was determined using an ROC curve and the maximum Cohen's kappa value as determined for each of the total scores.

Imperiale et al. created a risk index using points developed from the β-coefficients derived from a logistic regression analysis to detect advanced neoplasia.[44] There were 3,025 individuals in the development data set and 1,475 in the validation dataset. There were five factors investigated in the model and included age, gender, body fat, history of smoking cigarettes and familial history of colon cancer as the independent variables, while advanced neoplasia was the outcome variable. The investigators used an approach described by Sullivan et al. when developing their point system.[45] The β-coefficients were used to derive the points for each variable, which were further categorized into risk groups (very low, low, intermediate, high) once the points were determined. Comparisons of model parameters for the development and validation sets were made using the risk and likelihood ratios.

These studies all used regression analysis to create scoring tools for use in prediction of risk of some health related issue. Scoring tools were created using different methodologies and analyses. However, the ultimate goal was the same, the

creation of a tool to be used in the assessment of patients. The analytic techniques from these studies were used to map out an approach to creating the pilot-scoring tool for this dissertation.

Changing the Methodological Approach to Examining GME In-State Retention

From a methodological perspective, there are no previous reports in the GME literature that have utilized logistic regression with cross-validation to analyze in-state retention in order to create an instrument to identify graduates likely to practice within the state of GME training. This dissertation utilized these analytic methods to produce a novel pilot-scoring tool to identify graduates from a Michigan-based GME institution who are likely to practice in Michigan post-training. A tool like this could be beneficial to Michigan-based GME institutions, as well as hospital and state physician recruiters within the state. However, further research will be required to determine the universal applicability of this pilot-scoring tool and are outside the scope of this dissertation.

Chapter Summary

An overview of the literature related to retention of physicians within physician shortage/rural areas, as well as in-state retention of GME graduates was provided. Limitations related to the methodologies used to examine retention related to these areas, as well as what's lacking from the GME retention literature, were also discussed. Lastly, studies related to the development of scoring systems using multivariate regression and cross-validation techniques were explored.

CHAPTER III

RESEARCH DESIGN

The purpose of this study was to examine individual level characteristics of graduates from a single GME sponsoring institution to develop a tool to identify graduates that are likely to practice medicine in the same state. The state in this study is Michigan. This chapter describes the study procedures and analytic techniques used to address the following two study objectives:

1. To use logistic regression with cross-validation to examine the individual characteristics related to whether or not graduates who trained in a Michigan-based GME sponsoring institution practice medicine in Michigan.
2. Create a scoring tool based on the logistic regression with cross-validation to categorize GME program graduates into groups: likely to practice in Michigan or not likely to practice in Michigan.

Study Procedures

Study Sample Description

All graduates (n=1161) from the 18 GME training programs offered by Grand Rapids Medical Education Partners (GRMEP) from 2000 through 2014 were included in the initial review. Residents and fellows who graduated from one of the GRMEP training programs and were currently still in training (e.g., additional training, fellowship) at the time of data collection were excluded from the review. Transitional year and preliminary surgery residents who left GRMEP after their one-year of training in order to

enroll in another residency training program were also excluded from the study. The rationale for this is that these graduates had not completed their GME residency training. Data from the graduates from 2000 through 2014 were used to build a predictive model and create a pilot-scoring tool.

Rationale for Inclusion of Independent Variables in the Regression Model

Multiple studies and reports have stated that a connection to the state of GME training is related to in-state retention.[30-35] This includes being born in or attending high school or medical school in the state in which GME training was completed. Based on this information, the following variables were selected to reflect a tie to the state of Michigan (e.g., born in Michigan, obtaining a bachelor's degree in Michigan, attending medical school in Michigan). An additional potential tie to the state of Michigan is the length of time spent in GME training (time program).

Where a physician completed GME training has also been reported to be predictive of GME location. Multiple reports in the literature state that many GME graduates practice within the state where they completed their GME training.[8,32,33] Completion of GME training in this case means that graduates did not leave the state in order to undergo further GME training. For this dissertation completion of GME in Michigan is defined as the resident/fellow went into practice after graduating from a GRMEP GME program.

Marriage (ever married) could also be considered a potential predictor for whether or not a graduate practices in Michigan. For example, if a resident or fellow marries someone from Michigan during their training, they may be more likely to

practice in Michigan post-GME than if their spouse was from somewhere else. Conversely, if a resident or fellow is already married prior to starting their training, they may have ties established somewhere else and plan to go back after training. Spousal influence was shown to be an important factor in practice location decisions in multiple studies.[24,28,46]

Type of program (primary care vs. non-primary care) was also included for testing in the model. Graduating from a primary care program was reported to be associated with the practicing within the state of GME completion.[30,31,33,35] Female gender was also shown to be an important factor related to in-state retention after GME training.[30,31] The investigators of these studies reported that females were more likely to practice within the state of GME training than males. Therefore, gender was included as a study variable.

Lastly, temporary visa holders have restrictions as to how long they can be in the country, as well as on where they can practice after GME graduation. If these individuals want to remain in the US to practice, they are restricted to practicing in physician shortage areas.[35] Therefore, visa status was included in the model to assess whether this status had any influence (positively or negatively) on practicing within the state of Michigan.

Rationale for Exclusion of Variables from the Regression Model

Age at graduation was not accounted for in the logistic regression model. Since the outcome variable is whether or not the graduate has ever practiced in Michigan at any point in time post-training, age at the time of graduation does not have much

meaning since this decision could happen well beyond graduation. Also, more than half of the residents that graduate from the GME programs go on to do a fellowship after residency (lasting 1-3 years) prior to determining their practice location.

Time since graduation was also not accounted for in the logistic regression model. Time since graduation is a reasonable consideration, since one could propose that an individual who graduated in the year 2000 would have 10 more possible years to consider Michigan as a place to practice medicine than someone who graduated in 2010. However, inclusion of this variable in a scoring system of future fellowship and residency graduates is problematic, as time since graduation would have no meaning for someone who is being evaluated prior to graduation.

A sensitivity analysis was conducted to assess the relationship of time since graduation to the outcome variable, ever practiced in Michigan. A variable was created that included three time groupings. Graduates were grouped into those that graduated from 2000-2004, 2005-2009 and 2010-2014. This variable was included in a logistic regression model, along with the other predictor variables described above, to see if there was a significant relationship with ever practicing in Michigan. This variable was not a significant ($p>0.1$) predictor in the model, therefore, an assumption can be made that the time that a resident/fellow has to return to practice in Michigan is not confounding the dissertation results.

Data Collection

Table 3.1 shows each source, the type of documentation/information that could be obtained from the source, as well as the study variables related to the source. The

sources included the New Innovations database (Uniontown, OH), a residency data management system used by GME sponsoring institutions across the nation, GRMEP GME records, Google, the Department of Licensing and Regulatory Affairs (LARA) website, which gives access to licensing data of over a million individuals and businesses in Michigan and GRMEP program directors and coordinators.

Table 3.1

Data sources

Source	Information available	Data point
New Innovations	Residency applications, personal statements, CV, educational history, spouse information, visa status, program start/end dates, birth place	Gender, UGME Michigan, undergraduate Michigan, time in program, primary care program, birth place Michigan, visa status, ever married
GME department records	Paper applications, school transcripts, diplomas, documents related to visa status, last name change forms, job verifications post-training, CVs, location after graduation records	Gender, UGME Michigan, undergraduate Michigan, time in program, primary care program, birth place Michigan, visa status, completed GME training in Michigan
Google searches using graduates name	Physician bios with current employer, LinkedIn profiles	Gender, UGME Michigan, undergraduate Michigan, birth place Michigan, ever married, ever practiced in Michigan, completed GME training in Michigan
LARA licensing website	Michigan medical license history	Ever practiced in Michigan
GRMEP staff: GME Director, Program Director/Coordinator	Connection with former resident/fellow through email, LinkedIn, Facebook, and/or resident/fellow is GRMEP faculty	UGME Michigan, undergraduate Michigan, birth place Michigan, visa status, ever married, ever practiced in Michigan

The author used the different data sources in data collection efforts. The Director of GME was available to assist the author with data collection. In some

instances the author and the Director of GME reviewed applications together. In other cases, the author consulted the Director of GME, or other pertinent colleagues (e.g., program coordinators, program directors), when conflicting evidence was discovered and/or if a variable could not be found. For example, if undergraduate medical education could not be located in any of the data sources, the author reached out to the program coordinator to see if that information could be obtained through their office records. If data could not be located and/or if conflicts in the data could not be resolved, those data points were left blank.

The author started with a list of graduates from 6/1/2000 – 6/30/2014. Study numbers were assigned to each graduate and a correlation tool was created in order to de-identify the data. The number and name were then copied and pasted into an Excel document (correlation tool) and password protected. The names were then deleted from the Excel file with the study variables listed, leaving only the study number to be used as the identifier. The author used the correlation tool to identify individuals when needed. This way, the individual name was not stored with the study variables in order to protect the identity of the individuals included in the study.

The New Innovations database houses information related to current residents and fellows, as well as those that have graduated from GRMEP. Electronic documents housed in the database include residency applications, curriculum vitaes (CVs), and personal statements. This database was used to obtain information related to gender, undergraduate medical education, undergraduate education, program start date and graduation date, residency/fellowship program, place of birth, visa status, marital status

and location after graduation. Sections on the residency application and CVs related to hobbies and interests, as well as personal statements, were reviewed to search for information related to the study variables (e.g., wife, place of birth, educational information). Other sources from GME Department paper records included residency applications, CVs, personal statements, school transcripts, copies of diplomas, documents related to visa status, last name change forms, location of graduation documentation and job verifications post-training that could be used to gather information.

Google searches using the graduate's full name, if available, or just the first and last name, along with the type of medical degree (MD/DO), were also performed to locate missing data points. In most cases, Google searches would lead to the graduate's current employer where information about the graduate could be obtained. For example, Spectrum Health lists the physician's educational background (e.g., UGME institution, residency/fellowship location), as well as their age. If there was a physician biography available, the information was reviewed to obtain or verify information related to the study (e.g., birth place, educational background). For example, if the author was reviewing a physician bio online and educational history was listed, the data collected for this individual were checked against the website for consistency. Again, if there were discrepancies, colleagues were consulted to discuss. Any instances where there was doubt about the accuracy of the data, the variable was left blank.

The Department of Licensing and Regulatory Affairs (LARA) website was used to verify the outcome variable, ever practiced in Michigan. The GME Department

maintains records of location after graduation for all graduates. These data were verified by using LARA for all graduates to confirm whether a license to practice in Michigan was obtained after the graduation date.

Lastly, each residency/fellowship program has a program coordinator and a program director. The program coordinators keep the residency programs running smoothly and are very involved in the day-to-day activities of their residents. Also, many of the coordinators/directors keep in touch with their former residents/fellows through email, social media (e.g., LinkedIn, Facebook) or if the graduate works for the residency/fellowship program as a part of the teaching faculty. The program coordinators and directors were consulted during data collection to obtain data that the author could not locate or had questions about.

Data Preparation

Prior to data analysis, categorical study variables were coded (Table 3.2). Time in program was calculated using program start date and graduation date. Residency/fellowship programs were coded as primary care and non-primary care. Primary care includes Family Medicine, Pediatrics, Internal Medicine and Internal Medicine-Pediatrics. All other programs were designated as non-primary care and include all fellowship programs, as well as Radiology, General Surgery, Orthopaedic Surgery, Plastic Surgery, Surgical Critical Care, Vascular Surgery, and Obstetrics and Gynecology. Place of birth, location of undergraduate degree, location of medical education and completion on GME training were coded as Michigan/not Michigan.

Gender was coded as male/female and all other variables (married prior to/during residency, ever practiced in Michigan, visa status) were coded as yes/no.

Table 3.2

Coding for regression variables

Variables	Coding
Y = Practiced in Michigan (outcome variable)	0=No 1=Yes
X_1 = Gender	0=Male 1=Female
X_2 = Residency program type	0=Non-primary Care 1=Primary Care
X_3 = Place of birth	0=Not Michigan 1=Michigan
X_4 = Time in program	Years
X_5 = Ever married	0=No 1=Yes
X_6 = Location of Bachelor's degree	0=Not Michigan 1=Michigan
X_7 = Location of medical education	0=Not Michigan 1=Michigan
X_8 = Visa	0=No 1=Yes
X_9 = Location completed GME training	0=Not Michigan 1=Michigan

Missing Data

Missing data for each variable were assessed prior to the analyses. The percent of complete data for each variable are shown in Table 3.3. The majority of the variables were at least 90% complete.

Table 3.3

Percentage complete data for each study variable

Variable	% Complete n=988
Ever Michigan	988/988 (100%)
Primary Care	988/988 (100%)
UGME Michigan	988/988 (100%)
Gender	988/988 (100%)
Completed GME in MI	988/988 (100%)
Birth State MI	945/988 (95.6%)
Ever Married	925/988 (93.6%)
Visa Status	970/988 (98.2%)
Undergrad MI	988/988 (100%)
Time in Program	988/988 (100%)

When describing missing data, it is important to determine whether the absent values have a relationship to any of the other variables of interest. Allison, using the concepts described by Rubin, detailed three types of missing data: missing completely

at random (MCAR; there is no relationship between the missing data and any other variable), missing at random (MAR; missing data are due to a relationship with one of the other variables) and not missing at random (NMAR; missing data are due to the inherent nature of the variable itself).[47,48]

First, an analysis was performed to determine if missing data for the dissertation were MCAR. This was done using logistic regression to test whether there were any relationships between the missingness of any one variable and the other predictor variables in the model. For this dissertation, the outcome variable for this analysis was missing/not missing data for the ever married variable (the variable with the most missing data) and the predictor variables were gender, primary care, UGME Michigan, undergrad Michigan, birth state Michigan, visa numeric and completed GME in Michigan. If any variables were significantly related to the missingness of the ever married variable, the data would not be deemed to be MCAR. There were three statistically significant variables ($p<0.05$) in the regression model, gender, visa status and time in program. These results indicated that the missing data were not MCAR.

Next, the data were not considered NMAR since the cause of missingness was not related to the inherent nature of the variables themselves. In this case, the variables were missing due to lack of records to provide the information. Therefore the missing data for this dissertation were considered to be MAR.

Based on the decision that the missing data were MAR, the next decision was whether or not some form of imputation should be used to allow for the analysis of all of the subjects. Allison provides some insight into this issue, in his discussion of

imputation of categorical variables.[49] He developed a data set, then randomly eliminated 50% of the data for a categorical variable. He compared a complete case analysis (described as only using the half of the dataset that had complete data) against four different imputation techniques, at four different proportions for the categorical variable with missing data (0.5, 0.2, 0.1, and 0.01). The results showed that, for MAR data, all five methods had equivalent estimates of the β-coefficients, as well as equivalent standard deviations. His conclusion was that, for MAR data, there was no particular benefit to using an imputation technique. Based upon his findings, an imputation technique has not been used for this dissertation, instead, a complete case analysis was performed.

Logistic Regression Assumptions

Logistic regression analysis was used to address the first objective of the dissertation: To use logistic regression with cross-validation to examine the individual characteristics related to whether or not graduates who trained in a Michigan-based GME sponsoring institution practice medicine in Michigan. Prior to performing the analysis, data were checked to see if the assumptions for logistic regression were met using *Stata/IC 13.0 for Mac* (StataCorp, College Station, TX). The assumptions included (1) independence of errors, (2) all categories were mutually exclusive and exhaustive, (3) a linear relationship between the continuous predictor variables and the logit transformation of the dependent variable (practiced in Michigan Y/N) existed, (4) no multicollinearity of the predictor variables, and (5) no significant outliers or influential.

First, the assumption of independence was assumed. The data set did not include multiple observations for any one person, each row represented a unique individual. Also, all categories for each study variable were mutually exclusive and exhaustive.

Next, linearity between the continuous predictor variable of time in program and the logit transformation of the dependent variable (ever practiced in Michigan Y/N) was checked using the Box-Tidwell procedure.[50] This procedure consists of creating an interaction term between the natural log of the continuous variable and the original continuous variable to be tested in the model with the other variables. The interaction term for time in program was created using the natural log of the time in program variable and the original time in program variable. Next, the model was run including the interaction term. A significant (p<0.05) interaction term would indicate a non-linear relationship and would need further evaluation. Based on the assessment of the p-value for the interaction term, the assumption of linearity was met for time in program (p=0.997).

Next, multicollinearity among the predictor variables was assessed using the condition index and the regression coefficient variance-decomposition matrix. The condition index is derived from the eigenvalues, and represents collinearity related to the variable combinations. The condition index is the square root of the quotient of the largest eigenvalue divided by the smallest eigenvalue. Values indicative of multicollinearity suggested in the literature ranged from 10 – 30, it was decided to take

a conservative approach for this dissertation and go with a condition index >10 as suspect for collinearity.[51,52]

Table 3.4 shows the condition indices and their proportion of variance related to each regression coefficient. The condition index with a value of 16.92 was considered to be moderate to strong collinearity.[51] Based on this finding, the collinearity was further investigated by reviewing the variance decomposition proportions. Callaghan and Chen note that a high condition index (>10), along with two or more regression coefficient variances > 0.5, is indicative of problematic collinearity.[51] As shown in Table 3.3, both the constant and the time in program variable contribute greater than half of their variability to the eigenvalue associated with the condition index of 16.92. This indicated that a transformation to the time in program variable was warranted.

Snee and Marquardt suggested the use of centering in order to address the issue of collinearity.[53] For this study, the variable time in program was centered by creating a new variable (time in program – 3.5). The value of 3.5 was used since it was the mean time in program was 3.5 years. After centering the variable, the original time in program variable was removed from the model and replaced with the centered time variable. A reassessment of the condition index showed it had dropped from 16.92 to 8.48 (Table 3.5), indicating weak collinearity.[51] Based upon this finding, the centered time variable was used in further analyses.

Table 3.4

Multicollinearity assessment

Condition index	Constant	Primary care	UGME MI	Gender	Birth state	Ever married	Visa	Undergrad MI	Time program	GME completion MI
1.00	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00
1.98	0.00	0.01	0.04	0.01	0.04	0.00	0.15	0.03	0.00	0.00
2.89	0.00	0.07	0.02	0.03	0.03	0.05	0.43	0.02	0.01	0.00
3.32	0.00	0.04	0.00	0.70	0.00	0.07	0.04	0.00	0.00	0.00
3.92	0.00	0.70	0.01	0.15	0.00	0.00	0.35	0.01	0.00	0.00
4.87	0.00	0.00	0.48	0.00	0.77	0.00	0.00	0.02	0.00	0.00
5.32	0.01	0.02	0.03	0.09	0.01	0.77	0.02	0.03	0.03	0.11
5.66	0.00	0.04	0.23	0.01	0.09	0.00	0.00	0.49	0.03	0.29
5.97	0.01	0.03	0.19	0.00	0.06	0.06	0.00	0.40	0.08	0.28
16.92	**0.98**	0.09	0.00	0.00	0.00	0.04	0.01	0.00	**0.84**	0.30

Table 3.5

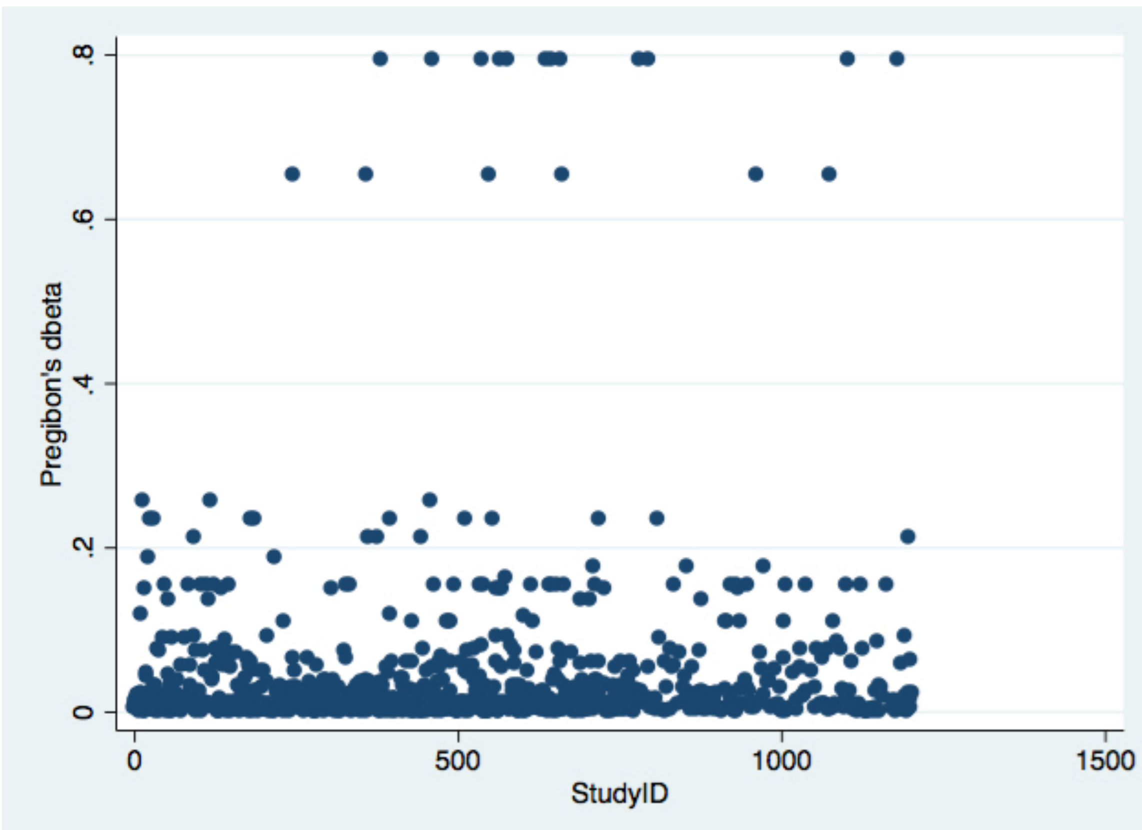
Multicollinearity reassessment

Condition index	Constant	Primary care	UGME MI	Gender	Birth state	Ever married	Visa	Undergrad MI	Time program	GME completion MI
1.00	0.00	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.00	0.00
1.82	0.00	0.02	0.03	0.00	0.03	0.00	0.14	0.02	0.06	0.00
2.28	0.00	0.01	0.01	0.02	0.01	0.02	0.00	0.01	0.57	0.00
2.87	0.00	0.01	0.02	0.14	0.02	0.02	0.59	0.02	0.09	0.01
3.19	0.01	0.02	0.00	0.61	0.00	0.16	0.01	0.00	0.02	0.01
3.82	0.00	0.87	0.01	0.10	0.00	0.01	0.22	0.00	0.09	0.00
4.55	0.00	0.00	0.48	0.00	0.76	0.00	0.00	0.02	0.00	0.00
5.13	0.03	0.04	0.03	0.10	0.00	0.56	0.02	0.02	0.02	0.37
5.37	0.00	0.00	0.41	0.01	0.16	0.01	0.00	0.89	0.01	0.02
8.48	0.95	0.03	0.01	0.01	0.00	0.21	0.01	0.02	0.14	0.58

The data were also checked for significant outliers and influential points by using Pregibon's delta beta (DBETA) influence measures.[54] This technique uses the DBETAs to determine influential observations, defined as data points which have a strong influence on the model performance.[54] DBETAS are a function of the standardized difference in betas with deletion of individual. The DBETAS were plotted to check for outliers (DBETA >0.25). Based on this criterion, there were 21 observations identified as outliers in the data set (Figure 3.1).

Figure 3.1

DBETA plot



The 21 outliers were removed to assess their influence on model performance. The difference between the performance of the models was minimal (Table 3.6). Table

3.7 shows the model parameters with outliers and Table 3.8 shows the model parameters with the outliers removed.

Table 3.6

Model statistics comparison with and without outliers

Model	# observations	Model χ^2	Prob > chi2
Model with outliers	879	281.65	<0.001
Model without outliers	858	284.67	<0.001

Table 3.7

Regression model statistics with outliers

Variable	β-coefficients	SE	*p* value	95% CI	
Primary Care	0.778	0.182	<0.001	0.422	1.134
UGME MI	0.762	0.262	0.004	0.249	1.275
Gender	-0.020	0.166	0.904	-0.346	0.306
Birth State MI	1.215	0.262	<0.001	0.701	1.729
Ever Married	0.583	0.180	0.001	0.231	0.936
Visa Status	-0.142	0.241	0.557	-0.614	0.331
Undergrad MI	1.021	0.264	<0.001	0.504	1.538
Time in Program (centered)	0.137	0.090	0.126	-0.039	0.314
Completed GME in MI	0.593	0.224	0.008	0.154	1.033

Table 3.8

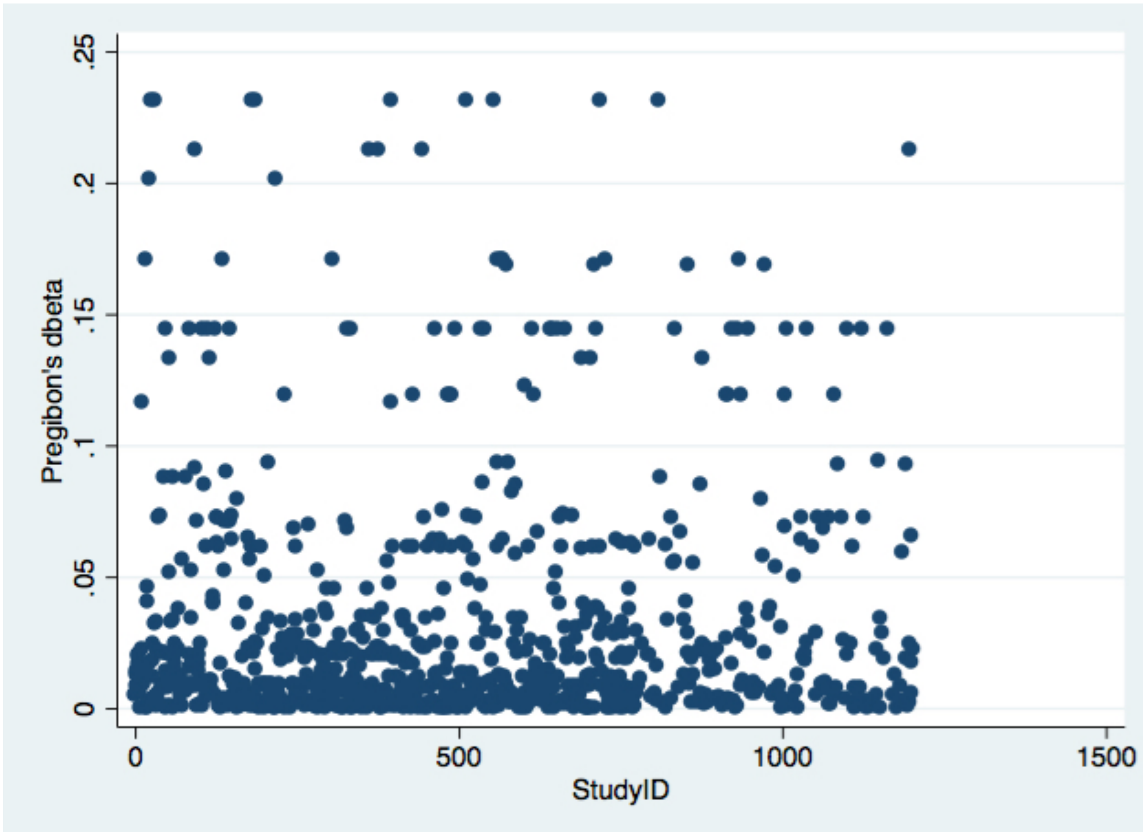
Regression model statistics without outliers

Variable	β-coefficients	SE	*p* value	95% CI	
Primary Care	0.797	0.183	<0.001	0.438	1.156
UGME MI	0.679	0.266	0.011	0.157	1.201
Gender	-0.003	0.170	0.988	-0.335	0.330
Birth State MI	1.275	0.266	<0.001	0.754	1.795
Ever Married	0.554	0.183	0.002	0.196	0.913
Visa Status	-0.161	0.258	0.534	-0.667	0.345
Undergrad MI	1.082	0.267	<0.001	0.559	1.604
Time in Program (centered)	0.137	0.090	0.129	-0.040	0.314
GME Completion in MI	0.605	0.226	0.007	0.162	1.048

The DBETAs were also examined without the outliers to assess whether there may be issues with other observations after their removal. No additional outliers (>0.25) were revealed a review of the DBETAS plot (Figure 3.2). Based on the performance of the models with and without outliers, it was determined to leave the outliers in the dataset.

Figure 3.2

DBETA plot after removing outliers



Analytic Procedures

This next section describes the analyses performed to address objective one of the dissertation: to use logistic regression with cross-validation to examine the individual characteristics related to whether or not graduates who trained in a Michigan-based GME sponsoring institution practice medicine in Michigan.

Logistic Regression

The first step was to perform a logistic regression using the best subsets logistic regression approach.[55,56] This analysis uses all different variable combinations of the

model to determine the best model subset. The criteria used in model evaluation included the adjusted R^2, Mallow's C, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). These criteria are suggested by Hosmer et al., as well as Lindsey & Sheather, to use in the selection of the optimal model when comparing multiple models.[55,56] The adjusted R^2 is used to determine the model that explains the most amount of variability within the outcome variable. Optimal values for Mallow's C have been described as either equal to the number of predictors plus one, or as the smallest value.[56,57] The AIC and BIC are both measures of model fit that are penalized for the number of parameters in the model.[56] The concern with the AIC is over-fitting and the concern for BIC is under-fitting.[58] The model with the lowest AIC and BIC is most desirable.[56]

The adjusted R^2, Mallow's C, AIC and BIC were used for model comparison and selection for this dissertation. In an ideal situation, all of these criteria would align on the same model. In cases where this does not occur, judgments must be made using these criteria. For example, Hosmer et al. note that the differences in these criteria can be so negligible (e.g., a difference <0.1) that the variables included in each model may need to also be considered.[55] Lindsey and Sheather also reported this issue.[56] In one example, they showed how the various criteria (adjusted R^2, Mallow's C, AIC and BIC) would indicate that either a 5, 6 or 8 factor model would be appropriate. They suggested that a 6 factor model would be a good compromise, as the values for the various criteria were only marginally different between the three possible models.

After model selection occurred, a logistic regression was performed using the reduced model. This model was assessed using the Hosmer & Lemeshow goodness of fit statistic and the AUC. The goodness of fit statistic is used to determine how well the model fits the data. A significant *p*-value (<0.05) would warrant further investigation of the model, while a *p*-value >0.05 represented good model-data fit.

The AUC, derived from an ROC curve analysis, was used to evaluate how well the model was discriminating between those that ever practiced in Michigan and those that never practiced in Michigan. Model performance based on the AUC was determined using the criteria established by Hosmer et al.[55] Based on these criteria, an AUC of 0.5 is no better than flipping a coin, while an AUC of 1.00 represents perfect discrimination. An AUC of $\geq$0.7 was considered acceptable and was the criterion used for this study.[55] Next, the model underwent two forms of cross-validation to check for over-fitting.

Over-Fitting

One of the concerns for any predictive model derived from a specific sample is whether or not it will still have value when applied to other data sets.[59,60] The primary concern is over-fitting. This occurs when the model is so tightly fitted to the data, that it may not perform the same when used with other data sets. The technique of cross-validation is a valuable tool to test the reproducibility of a data set, with a goal of testing the accuracy and validity of the model.[59] There are a variety of ways to do this, including hold-out methods, leave-one-out cross-validation, k-fold cross-validation and bootstrap techniques.

For example, the hold out method involves splitting the sample into two groups, a training set and a test set. The model parameters would be defined using the training set, and then evaluated by estimation of the error rate using the test set. There are two major concerns with this methodology. The first is that the sample may be too small to allow for splitting into two sub-sets. The second is that using a single test set and training set may lead to an unreproducible result, due to the spurious decision as to which data are in the test set and which are in the training set.[59]

One solution to this problem is to involve multiple samples of the entire data set, such as in a k-fold validation or bootstrapping technique. For the former, the data would be divided into k random samples. The latter uses random sampling with replacement. For this study, the model underwent both a five-fold and bootstrap internal cross-validation procedures to test for over-fitting.

Five-Fold Cross-Validation

The first method used for this dissertation was five-fold cross-validation. In this step, the data set was randomly split into five groups of similar size. Four of the groups were the test set, while the remaining group was used as the validation set. The test set was used to run the model and the validation set was used to determine how well the independent variables from the validation data set were predicting the dependent variable.

This process was repeated five times, so that each of the five groups of data were used as the validation set once, while all of the rest of the data were used as the test sets. The criterion for the analysis was the root mean square error (RMSE), as

derived from testing each of the five validation sets. The five RMSE values obtained were reviewed to determine if they were consistent with the RMSE obtained from the final logistic regression model. Consistency for this evaluation was defined as each of the five RMSEs from the cross-validation should be no greater than 15% different from the original RMSE.[59]

Bootstrap

The model also underwent bootstrapping to further check for over-fitting. The bootstrapping procedure uses random sampling with replacement. For example, as there were 800 people with complete data, the bootstrapping procedure used sampling with replacement to choose 800 individuals for the procedure. This random sample of 800 could contain person 23 five times, person 799 not selected at all, person 200 selected twice, etc., for a total of 800 random individuals. The bootstrapping procedure for this study was replicated 200 times.

The evaluation criterion for this analysis was the Harrell's C, which is equivalent to the AUC[33]. For the sake of consistency, the Harrell's C will be referred to as the AUC throughout the rest of the dissertation. The AUC derived for the original model was compared to the AUC from the bootstrap cross-validation procedure. The bootstrap method produced a corrective factor, which was used to create the over-fitting corrected estimate of the AUC.[40] The original and corrected values for the AUC were compared for consistency. For this step, consistency was defined that the average corrected AUC (derived from the 200 replications) should be no more than 15% different from the original AUC.[59]

Pilot-Scoring Tool

This section describes the methods used to address the second objective of the dissertation: Create a scoring tool based on the logistic regression with cross-validation to categorize GME program graduates into groups: likely to practice in Michigan or not likely to practice in Michigan.

The final model selected during the analyses related to objective one was used to create a scoring system to identify graduates who are most likely to practice in Michigan. The method of Sullivan et al. was used for this step.[45] The β-coefficients, obtained from the independent variables used in the final model, were used to create the scores. The β-coefficients were compared, with the lowest value representing the referent value, from which the remaining scores were determined. Each β-coefficient was divided by the lowest β-coefficient in order to determine the score for that variable. Products from this calculation were rounded to the nearest whole number.

Table 3.9 shows an example of Sullivan's methodology in the context of this dissertation.[45] In this example, scores were determined for primary care, undergraduate medical education in Michigan (UGME MI) and undergraduate education in Michigan (Undergrad MI). First, each variable's β-coefficient was multiplied by the values associated with each category (e.g., 0=no and 1-yes) for the variable (Table 3.9, column 2). Next, the products were divided by the lowest β-coefficient, which was associated with primary care (Table 3.9, column 3). Finally, the products of these calculations are rounded to the nearest whole number to create the final score associated with each variable (Table 3.9, column 4).

Table 3.9

Methodology for assigning scores to regression variables

Variable	β-coefficient	Score computation	Rounded score
Primary care	0.69		
Yes = 1	1*0.69=0.69	0.69/0.69=1	1
No = 0	0*0.69=0	0/0.69=0	0
UGME MI	1.26		
Yes = 1	1*1.26=1.26	1.26/0.69=1.82	2
No = 0	0*1.26=0	0/0.69=0	0
Undergrad MI	1.37		
Yes = 1	1*1.37=1.37	1.37/0.69=1.99	2
No = 0	0*1.37=0	0/0.69=0	0

The next step included assigning scores to all of the graduates with complete data for the variables included in the final model. Each coded variable in the model was multiplied by the score associated with it. The scores for each variable were then added together to determine the final score for individuals. When complete, each person had a score associated with whether or not the graduate ever practiced medicine in the state of Michigan. Next, the ROC analysis was performed to obtain the AUC to see how well the model was discriminating between graduates that practiced in Michigan and those that did not. For the ROC analysis, the x-axis represents the false positive rate (1-specificity) and the y-axis represents the true positive rate (sensitivity). Once again, the

AUC criteria of Hosmer et al. were used to evaluate the AUC.[55] An AUC $\geq$0.7 was considered to be acceptable.

Next, a score cut-point was determined in order to maximize the predictive model of the scoring system, which in this case means to maximize the ability to determine whether or not a graduate will ever practice in Michigan. Perkins and Schisterman suggested the use of the Youden Index, also known as Youden's J.[61] This has been shown to be a good estimate of the preferred score cut-point, and is equal to the maximum value of sensitivity plus specificity minus 1, as derived from all possible cut-points used to create the ROC curve. This preferred score cut-point represents the score best suited to discriminate between the two groups being evaluated. For this step, the sensitivity (true positive rate) and the specificity (true negative rate) for each cut-point associated with a score were determined from the ROC analysis.

Table 3.10 shows an example of the cut-points derived from a ROC analysis, with their associated sensitivity and specificity. A third column shows the values for the sensitivity plus specificity minus 1. The cut-point that maximizes the sensitivity (i.e., the ability to identify those who practice in Michigan) while maximizing the specificity (i.e., the ability to identify those who do not practice in Michigan) can be identified through examination in this table. For example, if a cut-point of >= 0 is selected, this can be interpreted as everyone would be identified as practicing in Michigan. We would have a 100% true positive rate, however, our true negative rate would be 0%. Conversely, if we chose >= 5 as the cut-point, this is the equivalent as stating that anyone with a score of five or above would be identified as staying in Michigan. In this case, our true positive

rate would be very low, while our true negative rate would be really high. Therefore, in this example, the Youden's J is 0.44 and is associated with a cut-point of a score greater than or equal to 2. Meaning, those with a score of 2 or more would be identified as practicing in Michigan and those with a score of 0 or 1 would be identified as not practicing in Michigan. The true positive rate is 60%, while the true negative rate is 84%.

Table 3.10

Cut-points derived from an ROC analysis

Score	Sensitivity	Specificity	(Sensitivity+Specificity) - 1
>=0	1.0	0.0	0.0
>=1	0.84	0.52	0.36
>=2	0.60	0.84	0.44
>=3	0.54	0.89	0.43
>=4	0.44	0.93	0.37
>=5	0.24	0.97	0.21

The final step in the procedure included creating a scoring tool. The tool created included questions based upon each study variable that was included in the final model, with points assigned to the response to each question. For example, if birth state in Michigan were to be included in the final model, the question in the tool could be "Was the resident/fellow born in Michigan?" The response options would be yes or no with

points assigned to each response (yes= 1 point and no=0 points). This would be done for each variable included in the model.

Chapter Summary

Described in this chapter were the objectives of the study and the methodology used to carry out the study. The study procedures, including a description of the sample, data points and their rationale for inclusion, data collection, and methods for checking the assumptions of logistic regression, were detailed. Lastly, the analytical procedures used to determine the variables for the pilot-scoring tool were described.

CHAPTER IV

RESULTS

Summary statistics are described for the data in this chapter. The results of the logistic regression and cross-validation analyses are also detailed. Next, the development and results of the performance of the scoring system for the data are discussed. Lastly, the scoring tool derived from the final regression model is presented.

Summary Data

Data for the analysis included 988 graduates. Summary data for the sample are shown in Table 4.1. Just over half of the graduates practiced in Michigan at some point after graduation. The sample consisted of mostly males. Close to a third of the sample attended UGME in Michigan, attended an undergraduate institution in Michigan or were born in Michigan. Just under half of the graduates were from a primary care program. Approximately 80% completed their GME training in Michigan.

Table 4.1

Summary data for the sample

Variable	
Ever Michigan	504/988 (51.1%)
Primary care	475/988 (48.1%)
UGME Michigan	302/988 (30.6%)
Gender (% Male)	573/988 (58%)
Completed GME in MI	793/988 (80.3%)
Birth state MI	266/945 (28.0%)
Ever married	661/925 (71.5%)
Visa status	126/970 (13%)
Undergrad MI	336/988 (34.0%)
Time in program (mean $\pm$SD)	3.5$\pm$1.0 yrs

Logistic Regression

The first analysis was performed to address objective one, which was to use logistic regression with cross-validation to examine the individual characteristics related to whether or not graduates who trained in a Michigan-based GME sponsoring institution practice medicine in Michigan, used a best subsets logistic regression approach. The outcome variable was ever practiced in Michigan and the predictor variables included birth state in Michigan, undergrad in Michigan, primary care, ever

married, UGME in Michigan, completed GME in Michigan, time in program, gender and visa status.

The best subsets analysis produced the 11 best variable combinations for the model (Table 4.2). This table includes the adjusted R^2, Mallow's C, the AIC and the BIC, which were used to determine the best model for further analyses. The criteria described in Chapter III for this evaluation did not converge on the same model. The model with the lowest BIC was a five variable model, which included birth state in Michigan, primary care, undergrad Michigan, UGME Michigan and ever married, though the Mallow's C was higher than the number of variables in the model. The seven variable model (birth state in Michigan, primary care, undergrad Michigan, UGME Michigan, completed GME in Michigan, ever married, program years) had the highest adjusted R^2, a Mallow's C less than the number of parameters in the model and lowest AIC.

The problem of selecting the model with the lowest BIC is the concern for under-fitting, while problem of selecting the seven variable model with the lowest AIC and maximum adjusted R^2 is the concern for over-fitting. With this in mind, the six variable model, which included birth state in Michigan, primary care, undergrad Michigan, UGME Michigan, completed GME in Michigan and ever married was selected. This model had a Mallow's C smaller than the number of parameters in the model coupled with an adjusted R^2, AIC and BIC similar to that of the five and seven parameter models.

Table 4.2

Results of the best subsets logistic regression analysis

Block	Model parameters	Adj. R^2	Mallow's C	AIC	BIC
1	BS MI	0.2051	85.229	1076.302	1085.860
2	BS MI, UG MI	0.2368	47.905	1041.547	1055.883
3	BS MI, PC, UG MI	0.2527	29.756	1024.101	1043.217
4	BS MI, PC, UG MI, Married	0.2625	18.814	1013.380	1037.274
5	BS MI, PC, UG MI, Married, UGME MI	0.2704	10.336	1004.944	**1033.617**
6	BS MI, PC, UG MI, Married, UGME MI, GME MI **(selected)**	0.2740	6.991	1001.573	1035.025
7	BS MI, PC, UG MI, Married, UGME MI, GME MI, PY	**0.2754**	6.358	**1000.915**	1039.145
8	BS MI, PC, UG MI, Married, UGME MI, GME MI, PY, Visa	0.2749	8.000	1002.553	1045.562
9	BS MI, PC, UG MI, Married, UGME MI, GME MI, PY, Visa, Gender	0.2740	10.000	1004.553	1052.341

BS MI=Birth State MI; UG MI=Undergrad MI; PC=Primary Care; GME MI= Completed GME in MI; PY=Program Years

The final model selected included 889 observations, see Table 4.3 for summary data for the final model sample.

Table 4.3

Summary data for sample included in the final model

Variable	%
Ever Michigan	458/889 (51.5%)
Primary care	425/889 (47.8%)
UGME Michigan	268/889 (30.2%)
Birth state MI	253/889 (28.5%)
Gender (% male)	507/889 (57.0%)
Ever married	631/889 (71.0%)
Undergrad MI	293/889 (33.0%)
Completed GME training in MI	715/889 (80.4%)

Table 4.4 shows the β-coefficients, standard errors, p-values and 95% confidence intervals for the variables included in the model. Other model parameters that were assessed included the Hosmer & Lemeshow goodness of fit statistic (χ^2=3.21; *p*=0.201) and the Nagelkerke pseudo R^2 =0.361, percentage correctly classified by the model (73.57%) and the sensitivity (true positive rate: 60.48%) and specificity (true negative rate: 87.47%). The non-significant goodness of fit statistic suggests that there is good model data fit. The pseudo R^2 implies the model explains 36.1% of the variability within

the outcome variable. The model correctly classified 73.6% of the observations in the data set as either practicing in Michigan or not practicing in Michigan.

Table 4.4

Final regression model statistics

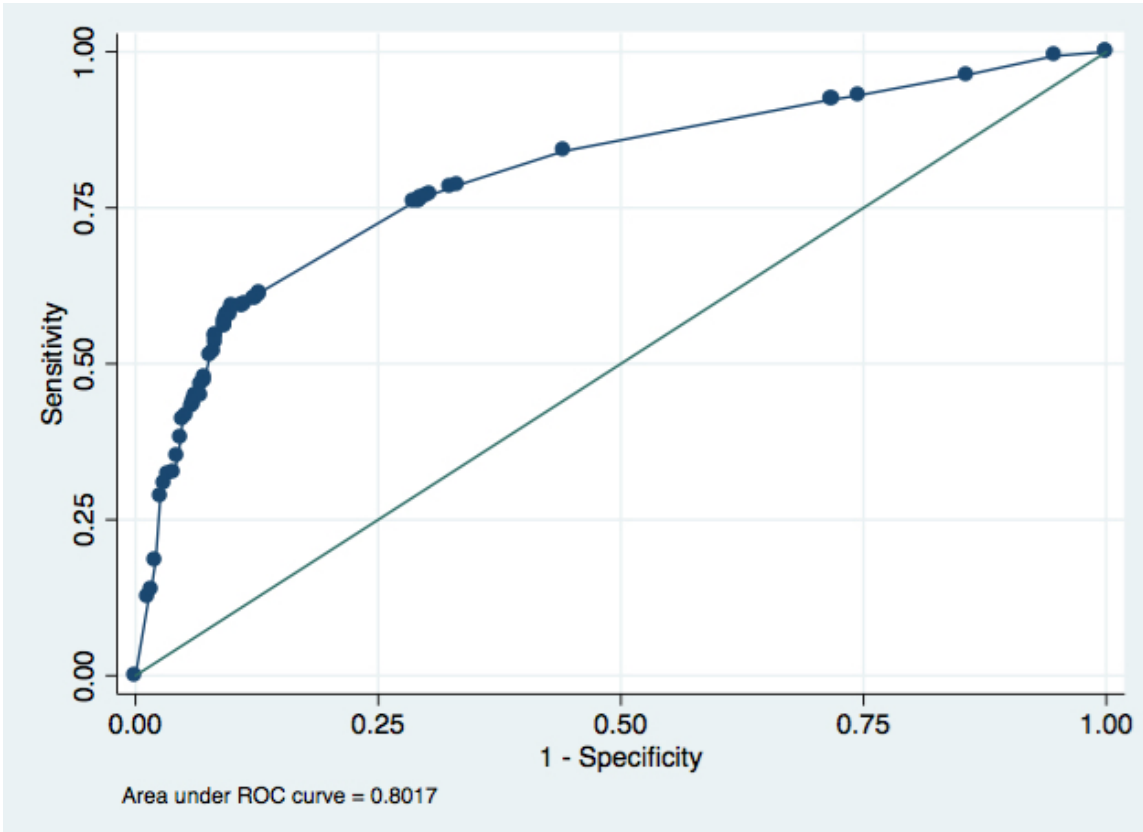
Variable	β-coefficient	SE	*p* value	95% CI	
Birth state MI	1.214	0.261	<0.001	0.703	1.725
Primary care	0.694	0.162	<0.001	0.376	1.012
Undergrad MI	1.038	0.262	<0.001	0.523	1.552
Ever married	0.629	0.178	<0.001	0.280	0.978
UGME MI	0.784	0.260	0.003	0.274	1.295
Completed GME in MI	0.474	0.208	0.022	0.067	0.881

The results of the ROC analysis are shown in Figure 4.1. The x-axis represents the false positive rate, while the y-axis shows the true positive rate. The diagonal line represents the value related to no discrimination, similar to flipping a coin. The line with the circular symbols represents the data for this study.

The AUC, derived from the ROC analysis, was 0.802. This final model met the criteria (AUC $\geq$0.7) established by Hosmer et al. deeming it an acceptable model for discriminating between those that ever practiced in Michigan and those that did not.[55]

Figure 4.1

ROC for final model



Five-Fold Cross-Validation

The next analysis addressed the second part of objective one, which was to use logistic regression with cross-validation to examine the individual characteristics related to whether or not graduates who trained in a Michigan-based GME sponsoring institution practice medicine in Michigan. A five-fold cross-validation procedure was performed using the selected model to assess for over-fitting the model to the data. In this step, the data set was randomly split into five groups of similar size. Four of the groups were the test set, while the remaining group was used as the validation set. The

test set was used to run the model and the validation set was used to determine how well the independent variables from the validation data set predicted the dependent variable.

This process was repeated five times, so that each of the five groups of data were used as the validation set once, while all of the rest of the data were used in the various test sets. The criterion for the analysis was the RMSE, as derived from testing each of the five validation sets. Then the five RMSE values obtained were checked for consistency against the RMSE from the final model. Consistency for this evaluation was defined as each of the five RMSE should be less than 15% different from the original RMSE.[59] The results of the RMSE comparison are shown in Table 4.5. Based on the evaluation criterion ($\leq$15% different) for this analysis, it does not appear that there are issues with over-fitting the model to the data.

Table 4.5

Five-fold cross-validation RMSE comparison

	RMSE
Original model	0.423
1	0.427
2	0.447
3	0.410
4	0.427
5	0.418

Bootstrap

A second form of cross-validation to address objective one was also performed. This step included performing the bootstrap technique to further check the final model for over-fitting. The bootstrap procedure used random sampling with replacement. This analysis was run on 889 graduates, therefore 200 random samples of 889 graduates were obtained. This random sample of 889 could contain graduate 757 five times, graduate 12 not selected at all, graduated 579 selected twice, etc., for a total of 889 subjects. The bootstrapping procedure for this study was replicated 200 times.

The evaluation criterion for this analysis was the AUC. The value for the AUC, was compared to the AUC from the bootstrap cross-validation procedure for consistency, defined as less than 15% different from one another. The original AUC and corrected AUC are 0.802 and 0.803, respectively. As described by Harrell, the corrected AUC represents a value which has been adjusted for bias allowing the calculation of an over-fitting corrected estimate.[40] Based on the evaluation criterion ($\leq$15% different) set for this analysis, over-fitting is not an issue.[59] Both the five-fold cross-validation method and the bootstrap method suggested that over-fitting was not a concern.

Pilot-Scoring Tool

This next section describes the results of the second objective of the dissertation, which was to use the results from the logistic regression with cross-validation analysis to create a scoring mechanism to categorize GME program graduates into groups: likely to practice in Michigan or not likely to practice in Michigan. The β-

coefficients associated with the predictors from the final model were used to create a scoring system to identify graduates who are most likely to practice in Michigan. The method of Sullivan et al. was used for this step.[45]

Table 4.6 shows how the scores were calculated for each variable included in the final model. This was done by first establishing the reference category for each variable. The first column of Table 4.6 shows the reference categories for each variable. In this case, all of the variables in the model are yes/no variables, with yes coded as one and no coded as zero.

The second column in Table 4.6 shows the β-coefficient associated with each variable and the result of the β-coefficient multiplied by each coded variable. For example, the β-coefficient for birth state Michigan is 1.214. This was multiplied by 1 corresponding with a yes response and 0 corresponding with a no response. Next, the lowest β-coefficient of all the variables was used to determine the value corresponding to one point on the scoring scale. In this case, the completed GME in Michigan variable had the lowest β-coefficient (0.474).

The third column of Table 4.6 shows the computation for the score. This involved dividing each of the products from the column two calculations by the lowest β-coefficient (completed GME in Michigan=0.474). For birth state Michigan, this involved dividing 1.214 by 0.474 for the yes category and 0 by 0.474 for the no category. The last column in Table 4.6 shows the results of the calculations in column three rounded to the nearest whole number. The scoring system ranged from 0-10 points. A

yes response to being born in Michigan was associated with three points, a yes response to graduating from a primary care program was associated with one point, a yes response to obtaining an undergraduate degree in Michigan corresponded with two points, a yes response to ever being married corresponded with one point, a yes response to completing UGME in Michigan corresponded with two points and one point would be given to anyone who had a yes response to completing GME in Michigan.

Table 4.6

Score calculation of predictor variables in the final regression model

Variable	β-coefficient	Score computation	Rounded score
Birth state MI	1.214		
Yes = 1	1*1.214=1.214	1.214/0.474=2.561	3
No = 0	0*1.214=0	0/0.474=0	0
Primary care	0.694		
Yes = 1	1*0.694=0.694	0.694/0.474=1.464	1
No = 0	0*0.694=0	0/0.474=0	0
Undergrad MI	1.038		
Yes = 1	1*1.038=1.038	1.038/0.474=2.189	2
No = 0	0*1.038=0	0/0.474=0	0
Ever married	0.629		
Yes = 1	1*0.629=0.629	0.629/0.474=1.327	1
No = 0	0*0.629=0	0/0.474=0	0
UGME MI	0.752		
Yes = 1	1*0.784=0.784	0.784/0.474=1.654	2
No = 0	0*0.784=0	0/0.474=0	0
Completed GME in MI	0.474		
Yes = 1	1*0.474=0.474	0.474/0.474=1.000	1
No = 0	0*0.474=0	0/0.474=0	0

The next step included assigning scores associated with whether or not the graduate ever practiced medicine in the state of Michigan to all of the graduates in the dataset with complete data. Each coded variable in the model was multiplied by the score associated with it. The scores for each variable were then added together to determine the final score for individuals. When complete, each person had a score associated with whether or not the graduate ever practiced medicine in the state of Michigan.

An ROC analysis using the scores derived from the model was then performed to obtain the optimal cut-point for the scoring tool (Figure 4.2). The sensitivity (true positive rate) and the specificity (true negative rate) for each of the score cut-points (0-10) are shown in Table 4.7. The cut-point that maximized the ability to determine whether or not a graduate ever practiced in Michigan was determined using Youden's J.[61] The percentage associated with Youden's J is equal to the maximum value of sensitivity and specificity minus 1, as derived from all possible cut-points used to create the ROC curve.

Table 4.7 shows that the Youden's J for this analysis is 0.486, which is associated with a cut-point of 4. This cut-point has a sensitivity of 61.6% (true positive rate), a specificity of 87.0% (true negative rate) and correct classification rate of 73.9%. The interpretation for this cut-point means that individuals with a score $\geq$4 are more likely to practice in Michigan, than those who receive a score <4.

Figure 4.2

ROC score cut-points

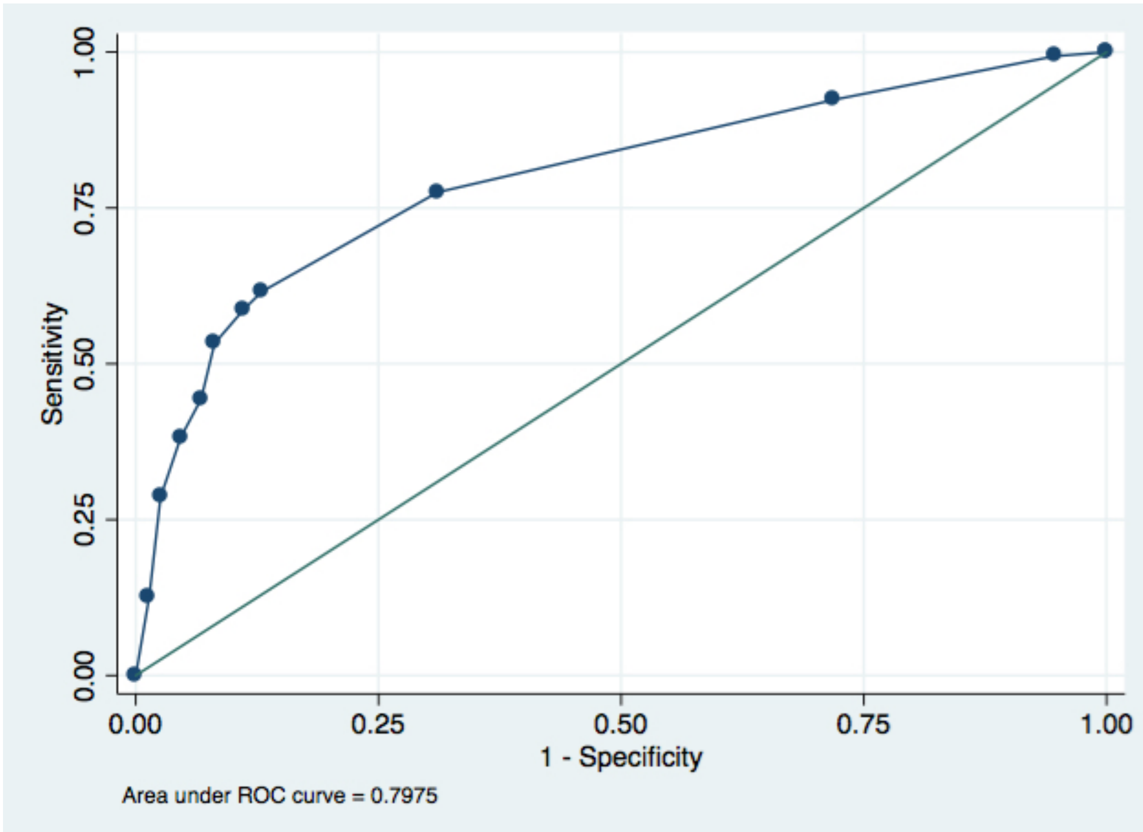


Table 4.7

Sensitivity and specificity for score cut-points

Score	Sensitivity	Specificity	Correctly classified	(Sensitivity+Specificity)-1
>=0	1.000	0.000	51.5%	0.000
>=1	0.993	0.053	53.8%	0.046
>=2	0.924	0.281	61.2%	0.205
>=3	0.775	0.689	73.3%	0.464
>=4	**0.616**	**0.870**	**73.9%**	**0.486**
>=5	0.587	0.889	73.3%	0.476
>=6	0.533	0.919	72.0%	0.452
>=7	0.443	0.933	68.1%	0.376
>=8	0.382	0.954	65.9%	0.336
>=9	0.286	0.975	62.0%	0.261
>=10	0.127	0.986	54.3%	0.113
>10	0.000	1.000	48.5%	0.000

The AUC for this analysis was 0.7975 and exceeded the criterion of acceptable discrimination $\geq$0.7 used for this dissertation.[55] Meaning the model was discriminating quite well between graduates who ever practiced in Michigan and those that did not.

See Appendix A for the pilot-scoring tool. The first column contains the questions created for each variable included in the final model the tool was derived

from, as well as the points associated with a yes or no response. The next column is for recording the points for each response to the questions. The last row of the tool is where the total score, calculated from the points recorded, will be recorded. Score interpretation is also shown. An electronic form that auto-calculates scores may also be an option for this tool.

Chapter Summary

Described in this chapter were the results of the analytical procedures used to create and cross-validate the model from which scoring system was derived. Also detailed were the results of the analyses used to assess the pilot-scoring tool for the 2000-2014 graduates. A description of the pilot-scoring tool derived from the logistic regression with cross-validation analyses was also discussed.

Chapter V

DISCUSSION

The purpose of the dissertation and study objectives are presented in this chapter. Next, a discussion of the results for to each objective is detailed. Lastly, limitations of the work, suggestions for future research and study conclusions are presented.

Study Purpose

Due to the overwhelming financial support of GME training by federal, state and training hospitals along with increased concerns over budgets and physician shortage issues, GME sponsoring institutions are under increased pressure to demonstrate ROI to funding sources. One suggested method for demonstrating ROI is to track in-state retention rates of graduates.[21,22] A mechanism for identifying residents/fellows during training who are likely to practice in the state in which they train could allow for targeted recruitment of these physicians, which could possibly result in higher in-state retention of GME graduates.

The purpose of this study was to examine individual level characteristics of graduates from a Michigan-based GME sponsoring institution, in order to develop a tool to identify graduates that are likely to practice medicine in the state in Michigan. The following objectives were used to guide the dissertation.

1. To use logistic regression with cross-validation to examine the individual characteristics related to whether or not graduates who trained in a Michigan-based GME sponsoring institution practice medicine in Michigan.
2. Create a scoring tool based on the logistic regression with cross-validation to categorize GME program graduates into groups: likely to practice in Michigan or not likely to practice in Michigan.

Research Objective One

The Michigan Connection

The regression analysis used to address the first objective produced a model that included six predictor variables: born in Michigan, medical school in Michigan, undergraduate education in Michigan, GME education in a primary care residency, completion of GME training in Michigan and whether or not the physician had married prior to or during residency. Based on the results of this dissertation, an individual who was born in Michigan and did both undergraduate and medical education in Michigan who completed residency/fellowship in a Michigan-based GME program is highly likely to practice in Michigan post-training. The three variables related to Michigan support the theory that GME graduates with some tie to the state may be more likely to practice in the state in which they trained than those with no tie to the state.[9,30-35,46]

Seifer et al. showed that physicians who attended medical school in the same state as GME training were almost four times as likely to be retained in the state to practice than those who did not.[31] The Georgia Statewide Area Health Education Center

stated that retention was just over 80% for physicians who graduated from Georgia based high schools, medical schools and residency programs.[9] Retention was just under 75% for those that graduated from Georgia based high schools and residencies. Bowman found similar findings for physicians in Virginia.[32] Many other reports have shown high in-state retention rates for GME graduates with a tie to their respective states.[30,34,35,46] Furthermore, for recruitment purposes, the Iowa medical society developed a database of physicians in residency that have a tie to the state of Iowa.[62] These include those that were born in the state, graduated from an Iowa-based medical school or are completing residency training in Iowa. The thought was that focusing on physicians with an Iowa connection would give a greater possibility of long-term retention of the physician within the state. Given these findings from previous reports, it is not surprising that the three variables with the largest β-coefficients (and the greatest weights in the pilot-scoring tool) were born in Michigan, UGME in Michigan and undergraduate in Michigan.

Final GME Training Location

The findings of this study suggest that completing GME training in Michigan, meaning the graduate did not leave the state for further training (e.g., elsewhere), was predictive of practicing in Michigan. This is consistent with much of the literature, which suggests that location at the end of GME training has been associated with a high percentage of graduates practicing within the state in which they trained.[8,32,33] In Bowman's study, a higher percentage of family medicine physicians that only completed GME training in Virginia practiced in Virginia post-training than physicians that only

attended medical school in Virginia and those that were born in Virginia, but did not receive their UGME or GME training there.[32] Fagan et al. reported that 57% of family physicians practiced within the same state as their residency training.[33] Physicians graduating from a GME program that are ready to go into practice as opposed to entering another GME program (e.g., fellowship) may be more likely to practice in the state in which they trained.

Primary Care and In-State Retention

Primary care was also included in the final model. The results showed that those who graduated from a primary care program were more likely to practice in Michigan than those that did not. This result is consistent with the study by Seifer et al. who reported that general practitioners were 1.4 times more likely to practice medicine within the state of GME training than specialists, which was a significant predictor in their regression model.[31] Armstrong et al. also reported that the need for primary care physicians in New York was greater than the that of graduates with more specialized training, which implies that there were more job opportunities available with in the state.[35] They supported this with additional data which showed that primary care physicians received an average of 4.3 job offers compared to 3 for specialists.

This finding for the current study may be driven by the need for primary care graduates in the state of Michigan and not as much of a need for the other graduates from different programs. In 2014, the state of Michigan had 293 locations, within its 83 counties, designated as primary care health professions shortage areas (HPSA).[63,64] At that time, only 63.6% of the primary care health needs were being met in those

designated areas and it was reported that it would take just over 200 physicians to alleviate the needs and to remove the shortage designations.[63] In 2015, location with the HPSA designation for primary care grew from 293 locations to 308 locations.[65]

With this growing need, job opportunities in Michigan may be more abundant for graduates from primary care programs. The majority of the primary care graduates make practice decisions immediately following graduation, whereas those in programs such as general surgery, radiology and orthopedics are more likely to go on to do more specialized training in a fellowship program. Therefore, primary care graduates may get recruited more heavily as graduation approaches, making them potentially more likely to practice in Michigan after graduation. Further, graduates who go on to do a fellowship in another state may be less likely to return if they have no connection to the state in which they did their training or no potential job prospects when they leave.

Marriage and In-State Retention

Whether or not the graduate was ever married was a significant predictor in the final regression model. In other words, graduates who were ever married were more likely to practice in Michigan than their single counterparts. Many studies have shown that spouses of physicians have an influence on practice location decisions.[24,28,46] Some possible explanation as to why married graduates are more likely to practice in Michigan are as follows. It may be that the resident/fellow met their spouse during their training and the spouse has ties (e.g., family, friends, career) to Michigan. Alternatively, the resident/fellow and their spouse have built a life during training and have established relationships that may influence the practice location decision to practice in

the state of GME training. Perhaps there were underlying financial incentives in the form of medical school loan repayment programs to practice in one of Michigan's rural areas that the graduate and their spouse thought was worth it to stay. Exploration of the underlying reason this variable was significant in the model is outside the scope of this study.

Variables Not in the Final Regression Model

Variables not included in the model were time in program, gender and visa status. It was interesting that time in program did not influence the decision to practice in Michigan. The theoretical rationale for including this variable was that residents who trained in the area for longer periods of time might be inclined to stay due to having more time to establish stronger ties to the community. However, the residents in those longer programs usually head off to do a fellowship somewhere else and may end up practicing where they complete their GME training. As noted earlier, there is a good evidence to indicate that physicians tend to practice medicine in the state in which they finish GME training.[8,32,33] Of the 211 GRMEP graduates who went on to fellowship training after graduation, 178 (84.4%) left the state to do so. Only 35% of those graduates returned to practice in Michigan after fellowship.

Additionally, this finding could be tied to the needs of certain specialties in Michigan. The primary care specialties are 3-4 year training programs, whereas the non-primary care specialties are 4-6 year training programs. If there are no jobs within the graduate's specialty, then the choice to stay or come back may not be any option for these graduates, making the length of time spent in the program immaterial. As the

need for primary care physicians is growing across the nation, the opportunities are greater for this group of graduates.[17]

The decision to practice in Michigan was not influenced by gender differences. Burfield et al. showed that 60% of female graduates practice in the state in which they underwent GME training, as opposed to only 50% of male graduates.[30] Similarly, Seifer et al. found that female GME graduates were 1.2 times more likely to practice medicine in the state of GME training than males, which was a statistically significant result on multivariate analysis.[31] These findings were absent in this dissertation. This could be due to the fact that far more females have entered into the physician workforce over the last 20 years.

Visa status was also not a significant predictor of practicing medicine in Michigan or not. Residents/fellows with temporary visas are only allowed to be in the country for a certain period of time before they must leave again. There are strict requirements for those on temporary visas regarding practicing in the US after graduation. If a visa holder wants to remain in the US, they must practice in a state or federally designated physician shortage location.[35] Michigan has many rural areas that hold an HPSA designation.[63] These types of opportunities, available both inside and outside of Michigan, may have influenced the fact that having a temporary visa does not have a positive or negative relationship with practicing in Michigan. Although not significant in the final model, 35.7% of the visa holders in the dissertation data set ended up practicing in Michigan at some point post-GME training.

Model Performance

The ROC analysis showed the model was able to discriminate between those that practiced in Michigan and those that did not just below the level of excellent discrimination (0.8), based on the criteria established by Hosmer et al.[55] The technique of cross-validation was used to test the reproducibility of a data set, with a goal of testing the accuracy and validity of the model. Over-fitting the model to the data is often a challenge in regression analyses thus the final model was also checked for over-fitting using two different methods.

The first was five-fold cross-validation. The criterion for excess over-fitting was a change in RMSE >15%, between the original RMSE and any of the five values derived from the cross-validation technique.[59] The maximum difference found in the dissertation data was 5.7%, which was much less than the criterion value.

The second procedure, bootstrapping, used sampling with replacement and is another method for checking the model data fit. For this analysis, the criterion for excess over-fitting was a change in AUC >15%.[59] The difference found in the dissertation data was 0.1%, which was far less than the criterion value. The analyses for the five-fold cross-validation and the bootstrap technique showed that over-fitting was not an issue for this model.

Research Objective Two

The second objective related to creating a pilot-scoring tool from the final model derived and tested using the logistic regression with cross-validation results. The six

predictors from the final model were used to create points related to each predictor and a cut-point for the total score was determined to use when deciding whether a graduate was likely to practice in Michigan or not. The methods of Sullivan et al. were used to develop the pilot-scoring tool.[45]

The points derived for each variable were birth state Michigan (3 points), UGME in Michigan (2 point), Undergrad Michigan (2 points), primary care (1 point), ever married (1 point) and completed GME training in Michigan (1 point). The ranges of scores for this tool were 0 points to 10 points. A cut-point of four was established using an ROC analysis, meaning that graduates with a score of $\geq$4 were likely to practice in Michigan and graduates with a score of 0-3 were not likely to practice in Michigan. A question for each variable was also developed for the final pilot-scoring tool for ease of use. This pilot-scoring tool can be used to assess residents/fellows at anytime during their training for identification for potential recruitment by hospitals, local offices/clinics and/or Michigan-based physician recruiters.

Contribution to Evaluation, Measurement and Research

Empirical evidence from a single Michigan-based GME sponsoring institution was used to examine variables related to whether or not a graduate practiced medicine in Michigan. Through the use of techniques from the field of research, including logistic regression with cross-validation, a novel pilot-scoring tool was developed to assist in the evaluation of whether or not GME graduates are likely to practice medicine in Michigan. The tool is a contribution to the field of evaluation. It provides a data driven approach to evaluating the likelihood a GME graduate would practice in Michigan. This tool could

be used to produce a list of identified graduates likely to practice in Michigan for hospitals and other physician recruiters in Michigan to use for targeted recruitment of these individuals.

Study Limitations

One limitation of the study was missing data. All of the variables in the data set were at least 90% complete. Since the data for this dissertation were considered to be MAR, the report of Allison would indicate that the data could be analyzed using the complete cases.[49] Additionally, as data set for this dissertation was large, missing data may not have as much of an influence on outcomes in this study as it would if the sample size were much smaller.[60] However, missing data can affect the outcome of analyses and therefore caution in the interpretation of results may be warranted.

Another limitation includes using data from a single institution, which can limit the generalizability of the findings. The pilot-scoring tool was created using data from one institution, therefore, this tool may not perform the same or include the same variables if data from multiple institutions in Michigan and/or other states were used to create the tool.

The extended study timeframe, 14 years, introduces issues related to factors that could influence practice location decisions. For example, Michigan went through a recession not that long ago, which could have potentially influenced the decision to practice in Michigan. Also, the employment opportunities are unlikely to be the same from one year to the next. The findings of this dissertation could also be influenced by those residents who graduated, but were still in fellowship training at the time of data

collection and therefore were not included in the study. Lastly, the more recent graduates (e.g., 2012, 2013, 2014) haven't had as much time to return to practice in Michigan as some of the early years (e.g., 2000, 2001, 2002). However, a sensitivity analysis, using a multivariate logistic regression, was conducted to determine the relationship of time since graduation and the outcome variable, ever practiced in Michigan, prior to the planned analyses for this dissertation. There was no significant effect seen for time since graduation.

Many factors remain unaccounted for in the model and subsequently the scoring tool. The model in this study accounted for 36% of the variability within the outcome variable. However, the data for the study were derived from data that were available through historical records, which leaves many variables that may further explain practice location decisions out of the model. These could include factors such as the weather, proximity to family, employment opportunities at the time of graduation, salary and benefit influences, military/loan repayment program status, spouses career/education.[35] Additionally, data from surveys and/or interviews of former residents could have been used to identify other influential factors related to practice location. However, both of these methods would prove time-consuming and would most likely further reduce the sample size.

Lastly, the quality of physicians that chose Michigan as a practice location and how long they practiced in the state are unknowns. A high in-state retention of mediocre physicians would not be a desirable outcome. Also, a high in-state retention of physicians who stay less than a year is an undesirable outcome. The addition of a

means to identify the quality of the graduates along with the scores of who is likely to practice in Michigan would further improve recruitment efforts and potentially the ROI to funders of GME training.   Tracking the length of time that a physician practices in the state of GME training is another addition descriptor that could be added to the performance marker of in-state retention.

Future Research

Further research could include pilot-testing the scoring tool on future graduates from GRMEP. Targeted recruitment of individuals identified as likely to stay and practice in Michigan could be undertaken. In-state retention rates could be tracked over time to determine if the targeted recruitment was worth the effort. For example, graduates could be scored early on in their last year of training. A list of randomly selected individuals likely to practice in Michigan could be given to hospital physician recruiters, as well as other Michigan-based physician recruiters in order to target those individuals to fill positions locally, regionally or statewide. In-state retention of the individuals likely to practice in Michigan on the list and those likely to practice in Michigan not on the list could be tracked over time to see if in-state retention rates are different between these two groups. If not, maybe targeted recruitment efforts are not necessary in this group and should potentially be focused at those not likely to stay. Similar studies using this design could include other institutions in Michigan and/or other institutions in the Midwest.

Another potential twist on this design would be to include random samples of those likely to practice in Michigan and those not likely to practice in Michigan on the

list for targeted recruitment. Comparisons of in-state retention rates between the groups could be used to provide supporting evidence, or not, of the scoring tools ability to discriminate between those likely to practice in Michigan or not.

The pilot-scoring tool could also be used to screen residency applicants (i.e., medical school graduates) for determining those that are likely to practice in Michigan after graduation. Applicants identified as having a score of four or higher prior to starting residency may be more likely to practice in Michigan post-training. This information could be used in the residency candidate selection process.

For example, the Family Medicine Program Director could score each residency candidate after reviewing his or her residency application and interviewing the candidate. When determining the rank of candidates for the program that rate similarly on the program's selection criteria (e.g., educational performance, personalities, volunteer experience) the score could be used to rank candidates more likely to practice in Michigan higher than those not as likely. In-state retention rates could then be tracked over time to see if there is an increasing trend over time. This could be done for five years and the in-state retention rates of the two graduating classes prior to implementation of this practice could be compared to the two graduating classes after implementation.

One last suggestion could be to assess the reliability of the predictability of the scoring tool. For example, scores could be generated for residents/fellows in the next two graduating classes (2016 and 2017). The practice location of these physicians would then be tracked through 2021. The data for the variable ever practiced in Michigan

could then be compared using ROC and AUC analyses to the original data for this dissertation.

Conclusion

The efforts of this dissertation produced a novel pilot-scoring tool for use in identifying GME graduates who are likely to practice medicine in Michigan post-training. Targeted recruitment of identified individuals may lead to increased in-state retention rates, which could translate into a means of demonstrating ROI to GME funding sources, particularly state and local sources.

REFERENCES

1. Association of American Medical Colleges. What Does Medicare Have to Do with Graduate Medical Education? aamc.org. https://www.aamc.org/advocacy/campaigns_and_coalitions/gmefunding/factsheets/253372/medicare-gme.html. Accessed October 23, 2015.
2. Association of American Medical Colleges. Why Teaching Hospitals Are Important to All Americans. aamc.org. https://www.aamc.org/advocacy/campaigns_and_coalitions/gmefunding/factsheets/253374/teaching-hospitals.html. Accessed October 23, 2015.
3. American Hospital Association. Fast Facts on US Hospitals. 2013 Annual Survey Data. http://www.aha.org/research/rc/stat-studies/fast-facts.shtml. Accessed October 23, 2015.
4. American Medical News. Uncap GME funding to break a dangerous training bottleneck. amednews.com http://www.amednews.com/article/20130204/opinion/130209978/4/. Accessed October 23, 2015.
5. Association of American Medical Colleges. 2013 State Physician Workforce Data Book. aamc.org https://www.aamc.org/download/362168/data/2013statephysicianworkforcedatabook.pdf. Accessed October 23, 2015.
6. Henderson, TM. Medicaid Graduate Medical Education Payments: A 50-State Survey. Association of American Medical Colleges.

https://members.aamc.org/eweb/upload/Medicaid%20Graduate%20Medical%20Education%20Payments%20A%2050-State%20Survey.pdf. Accessed October 23, 2015.

7. Accreditation Council for Graduate Medical Education. Report of ACGME Accreditated Sponsoring Institutions in Michigan. acgme.org. https://apps.acgme.org/ads/Public/Reports/Report/16. Accessed October 23, 2015.
8. French, R. Michigan gets med-school boom, doctor bust. *Bridge Magazine*. 2012:1-6. http://bridgemi.com/2012/06/michigan-gets-med-school-boom-doctor-bust/. Accessed October 23, 2015.
9. Georgia Statewide Area Health Care Network. 2013 Primary Care Summit Summary. Georgia Regents University website. http://www.gru.edu/ahec/documents/fy15gme_request.pdf. Accessed October 23, 2015.
10. Berwick D, Wilensky GR, Alexander B, et al. Graduate Medical Education That Meets the Nation's Health Needs. Institute of Medicine. http://www.iom.edu/Reports/2014/Graduate-Medical-Education-That-Meets-the-Nations-Health-Needs.aspx. Accessed October 23, 2015.
11. Metzler I, Ganjawalla K, Kaups K, Meara JG. The Critical State of Graduate Medical Education Funding. Bulletin of the American College of Surgeons. http://bulletin.facs.org/2012/11/critical-state-of-gme-funding/. Accessed October 23, 2015.

12. O'Shea JS. Reforming Graduate Medical Education in the US. *Backgrounder.* 2014;(2983): 1-14. Retrieved from http://report.heritage.org/bg2983. Accessed October 23, 2015.
13. Connolly B, DeVries J, Gifford B; Workgroup on Graduate Medical Education Funding. Report on Workgroup Efforts on GME Funding. michigan.gov website. https://www.michigan.gov/documents/mdch/18463_04_01_12_379998_7.pdf Accessed October 23, 2015.
14. Greene J. Snyder's Proposed State Budget Cuts Rural Hospitals, Graduate Medical Education Funding. Crain's Detroit Business. http://www.crainsdetroit.com/article/20150211/NEWS/150219947/snyders-proposed-state-budget-cuts-rural-hospitals-graduate-medical. Accessed October 23, 2015.
15. Michigan Primary Care Association. Conference Committee Finalizes FY 2016 Michigan Department of Community Health Budget. mpca.net. http://www.mpca.net/news/234177/Conference-Committee-Finalizes-FY-2016-Michigan-Department-of-Community-Health-Budget.htm. Accessed October 23, 2015.
16. Association of American Medical Colleges. Preserve Medicare Support for Physician Training. aamc.org. https://www.aamc.org/download/262676/data/gmefactsheet.pdf. Accessed October 23, 2015.

17. Association of American Medical Colleges. Physician Shortages to Worsen Without Increases in Residency Training. Kaiser Health News. https://kaiserhealthnews.files.wordpress.com/2011/01/md-shortage.pdf. Accessed October 23, 2015.

18. Meehan, M. Hospitals suffering under ACA, new Kentucky Hospital Association report says. kentucky.com http://www.kentucky.com/2015/05/08/3842438_hospitals-suffering-under-aca.html?rh=1. Accessed October 23, 2015.

19. Center for Health Workforce Studies. Michigan Physician Supply and Demand Through 2020. Institutue for Health Policy website. http://www.ihcs.msu.edu/documents/MichiganPhysicianSupplyandDemandthrough2020-Final.pdf. Accessed October 23, 2015.

20. Spero JC, Fraher EP, Ricketts TC, Rockey PH. GME in the United States: A review of State Initiatives. Cecil G. Sheps Center for Health Services Research, The University of North Carolina at Chapel Hill. https://www.shepscenter.unc.edu/product/gme-in-the-united-states-a-review-of-state-initiatives/. Accessed October 23, 2015.

21. Tripp Umbach. Expanding Graduate Medical Education in Indiana. Indiana University School of Medicine website. http://evansville.medicine.iu.edu/files/9314/3215/2212/State_of_Indiana_GME_Expansion_Approach_with_Southwest_IN_as_Pilot.pdf. Accessed October 23, 2015.

22. Commonwealth of Massachusetts Executive Office of Health and Humans Services. Report of the Special Commission on Graduate Medical Education. mass.gov website. http://www.mass.gov/eohhs/docs/eohhs/gme-final-report.pdf. Accessed October 23, 2015.
23. Rabinowitz HK, Diamond JJ, Markham FW, Paynter NP. Critical Factors for Designing Programs to Increase the Supply and Retention of Rural Primary Care Physicians. *JAMA*. 2001;286(9): 1041-1048.
24. Chan BT, Degani N, Crichton T, et al. Factors Influencing Family Physicians to Enter Rural Practice. *Canadian Family Physician*, 2005;51: 1246-1247.
25. Pellegrin KL. A Brief Survey to Identify Priorities for Improving Clinician Recruitment and Retention: Results from Hawaii Island Physicians. *Hawaii Journal of Medicine & Public Health*. 2012;71(4 Suppl 1): 45-49.
26. Robert Graham Center. What Influences Medical Student and Resident Choices? Robert Graham Center website. http://www.graham-center.org/dam/rgc/documents/publications-reports/monographs-books/Specialty-geography-compressed.pdf. Accessed October 23, 2015.
27. Hancock C, Steinbach A, Nesbitt TS, Adler SR, Auerswald CL. Why Doctors Choose Small Towns: A Developmental Model of Rural Physician Recruitment and Retention. *Social Science & Medicine*. 2009;69(9): 1368-1376.
28. Mayo E, Mathews M. Spousal Perspectives on Factors Influencing Recruitment and Retention of Rural Family Physicians. *Can J Rural Med*. 2006;*11*(4): 271-276.

29. Cameron P J. Professional, Personal and Community: 3 Domains of Physician Retention in Rural Communities. *Can J Rural Med*. 2012;17(2): 47-54.

30. Burfield WB, Hough DE, Marder WD. Location of Medical Education and Choice of Location of Practice. *J Med Educ.* 1986;61: 545-559.

31. Seifer SD, Vranizan K, Grumbach K. Graduate Medical Education and Physician Practice Location: Implications for Physician Workforce Policy. *JAMA*. 1995;274(9): 685-691.

32. Owen JA, Hayden GF, Bowman RC. Influence of Places of Birth, Medical Education, and Residency Training on the Eventual Practice Locations of Family Physicians: Recent Experience in Virginia. *Southern Medical Journal*. 2005;*98*(6): 674-674.

33. Fagan EB, Gibbons C, Finnegan SC, et al. Family Medicine Graduate Proximity to Their Site of Training: Policy Options for Improving the Distribution of Primary Care. *Access Fam Med.* 2015;47(2): 124-30.

34. Office for Healthcare Workforce Analysis & Planning. Retaining Physicians Educated in South Carolina Data Brief. Office for Healthcare Workforce Analysis & Planning website. http://www.officeforhealthcareworkforce.org/docs/Retaining%20Physicians%20Educated%20in%20South%20Carolina.pdf. Accessed October 23, 2015.

35. Armstrong D, Chung R, Forte GJ. 2014 New York Residency Training Outcomes A Summary of Responses to the 2014 New York Resident Exit Survey. Center for Health Workforce Studies website.

http://chws.albany.edu/archive/uploads/2015/04/2014_NY_Residency_Training_Outcomes.pdf. Accessed October 23, 2015.

36. How the Data Elements on the AMA Physician Masterfile Are Collected, Maintained and Verified. PDF filler website. http://www.pdffiller.com/51185066-primarysourcepdf-How-the-Data-Elements-on-the-AMA-Physician-Masterfile-are--Various-Fillable-Forms. Accessed October 23, 2015.
37. American Medical Association. AMA Physician Masterfile. American Medical Association website. http://www.ama-assn.org/ama/pub/about-ama/physician-data-resources/physician-masterfile.page. Accessed October 23, 2015.
38. Konrad TR, Slifkin RT, Stevens C, Miller J. Using the American Medical Association Physician Masterfile to Measure Physician Supply in Small Towns. *The Journal of Rural Health*. 2000;16(2): 162-167.
39. Sun GW, Shook TL, Kay GL. (1996). Inappropriate Use of Bivariable Analysis to Screen Risk Factors for Use in Multivariable Analysis. *Journal of Clinical Epidemiology*. 1996;49(8): 907-916.
40. Harrell Jr FE. Regression Modeling Strategies With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York, Springer-Verlag; 2001.
41. Dillman DA, Smyth JD, Christian LM. Internet, Mail and Mixed-Mode Surveys The Tailored Design Method (3rd ed.). Hoboken, NJ: John Wiley & Sons, Inc; 2009.
42. Hough CL, Caldwell ES, Cox CE, et al. Development and Validation of a Mortality Prediction Model for Patients Receiving 14 Days of Mechanical Ventilation. *Critical Care Medicine*. 2015;43(11): 2339-2345.

43. Zodpey SP, Deshpande SG, Ughade SN, Kulkarni SW, Shrikhande SN, Hinge AV. A prediction model for moderate or severe dehydration in children with diarrhoea. *J Diarrhoeal Dis Res.* 1999 Mar;17(1): 10-16.

44. Imperiale TF, Monahan PO, Stump TE, Glowinski EA, Ransohoff DF. Derivation and Validation of a Scoring System to Stratify Risk for Advanced Colorectal Neoplasia in Asymptomatic Adults: A Cross-sectional Study. *Ann Intern Med.* 2015;163(5): 339-346.

45. Sullivan LM, Massaro MM, D'Agostino RB. Presentation of Multivariate Data for Clinical Use: The Framingham Risk Score Functions. *Stat Med*. 2004;23(10): 1631-1660.

46. Rosenthal TC, Rosenthal GL, Lucas CA. Factors in the Physician Practice Location Puzzle: A Survey of New York State Residency-Trained Family Physicians. *The Journal of the American Board of Family Practice*. 1992;5(3): 265-273.

47. Allison PD. Handling Missing Data by Maximum Likelihood. SAS Global Forum 2012. Statistical Horizons website. http://statisticalhorizons.com/wp-content/uploads/MissingDataByML.pdf. Accessed October 23, 2015.

48. Rubin DB. Inference and Missing Data. *Biometrika*, 1976;63(3): 581-592.

49. Allison PD. Imputation of categorical variables with PROC MI. In SAS Users Group International, 30th Meeting (SUGI 30). SAS website. http://www2.sas.com/proceedings/sugi30/113-30.pdf. Accessed October 23, 2015.

50. Hilbe JM. Logistic Regression Models. CRC Press; 2009.

51. Callaghan KJ, Chen J. (2008). Revisiting the Collinear Data Problem: An Assessment of Estimator 'Ill-Conditioning' in Linear Regression. *Practical Assessment Research & Evaluation.* 2008;13(5): 1-6. http://pareonline.net/getvn.asp?v=13&n=5. Accessed October 23, 2015.

52. Belsley DA. A Guide to Using the Collinearity Diagnostics. *Computer Science in Economics and Management*. 1991;4(1): 33-50.

53. Snee RD, Marquardt DW. (1984). Collinearity Diagnostics Depend on the Domain of Prediction, the Model, and the Data. *The American Statistician*. 1984;38(2): 83-87.

54. Long JS, Freese J. Regression models for categorical dependent variables using Stata. Stata Press; 2006.

55. Hosmer DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression (3rd ed.). Hoboken, NJ: John Wiley & Sons, Inc; 2013.

56. Lindsey C, Sheather S. Variable Selection in Linear Regression. *Stata Journal*. 2010;10(4): 650.

57. Hocking RR. A Biometrics invited paper. The Analysis and Selection of Variables in Linear Regression. *Biometrics*. 1976;32(1): 1-49.

58. Burnham KP, Anderson DR. Multimodel Inference Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*. 2004;33(2): 261-304.

59. Harrell FEH, Lee KL, Mark DB. Tutorial in Biostatistics: Multivariable Prognostic Models: Issues in Developing Models, Evaluating Assumptions and Measuring and Reducing Error. *Statistics in Medicine*, 1996;15: 361–387.

60. Tabachnick BG, Fidell LS. Using Multivariate Statistics (6th ed.). Upper Saddle River, NJ: Pearson Education, Inc; 2013.
61. Perkins NJ, Schisterman EF. The Inconsistency of “Optimal” Cutpoints Obtained using Two Criteria Based on the Receiver Operating Characteristic Curve. *American Journal of Epidemiology.* 2006;163(7): 670–675.
62. Krupa C. How States are Keeping Doctors from Moving Out. American Medical News website. http://www.amednews.com/article/20111219/profession/312199944/1/. Accessed October 23, 2015.
63. The Henry J. Kaiser Family Foundation. Primary Care Health Professional Shortage Areas (HPSAs). The Henry J. Kaiser Family Foundation website. http://kff.org/other/state-indicator/primary-care-health-professional-shortage-areas-hpsas/. Accessed October 23, 2015.
64. Michigan Counties List. Michigan.gov website. http://www.michigan.gov/som/0,1607,7-192-29701_31713_31714-97053--,00.html. Accessed October 23, 2015.
65. Health Resources and Services Administration. Shortage Areas. HRSA Data Warehouse website. http://datawarehouse.hrsa.gov/Tools/DataByGeographyResults.aspx?geoTyp=State&geoCd=26. Accessed October 23, 2015.

Appendix A

Pilot-Scoring Tool

A. Pilot-Scoring Tool

	Points
Was the resident/fellow born in Michigan? *Yes = 3 points; No = 0 points*	
Will the resident/fellow graduate from a primary care program? *Yes = 1 points; No = 0 points*	
Did the resident/fellow receive a bachelor's degree from a Michigan-based college? *Yes = 2 points; No = 0 points*	
Was the resident/fellow ever married? *Yes = 1 points; No = 0 points*	
Did the resident/fellow graduate from a Michigan-based medical school? *Yes = 2 points; No = 0 points*	
Did the resident/fellow complete their GME training in Michigan? *Yes = 1 points; No = 0 points*	
Total Score	

Scores $\geq$4 = likely to practice in Michigan; Scores <4 = less likely to practice in Michigan

Appendix B

Human Subjects Institutional Review Board Letter

WESTERN MICHIGAN UNIVERSITY

Human Subjects Institutional Review Board

Date: May 12, 2015

To: Jessaca Spybrook, Principal Investigator
Tracy Frieswyk, Student Investigator for dissertation

From: Amy Naugle, Ph.D., Chair Amy Naugle

Re: HSIRB Project Number 15-05-13

This letter will serve as confirmation that your research project titled "GME Graduate Retention Rates: A Single Institution Study" has been **approved** under the **exempt** category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note: This research may **only** be conducted exactly in the form it was approved. You must seek specific board approval for any changes in this project (e.g., ***you must request a post approval change to enroll subjects beyond the number stated in your application under "Number of subjects you want to complete the study***)." Failure to obtain approval for changes will result in a protocol deviation. In addition, if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

Reapproval of the project is required if it extends beyond the termination date stated below.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: May 11, 2016

1903 W. Michigan Ave., Kalamazoo, MI 49008-5456
PHONE: (269) 387-8293 FAX: (269) 387-8276
CAMPUS SITE: 251 W. Walwood Hall