2005

Re-Engineering Graduate Medical Education: An Analysis of the Contribution of Residents to Teaching Hospitals Utilizing a Model of an Internal Medicine Residency Program

Ian M. Elius
University of South Florida

Re-Engineering Graduate Medical Education:

An Analysis of the Contribution of Residents to Teaching Hospitals Utilizing a Model of an Internal Medicine Residency Program

by

Ian M. Elius

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Industrial Engineering
Department of Industrial and Management Systems Engineering
College of Engineering
University of South Florida

Co-Major Professor: José L. Zayas-Castro, Ph.D.
Co-Major Professor: Peter J. Fabri, M.D.
Michael X. Weng, Ph.D.

Date of Approval:
July 12, 2005

Keywords: GME, healthcare system, linear programming, skill mix, proficiency

DEDICATION

To my mother, Dillia L.T. Elius, and the other members of my family.

ACKNOWLEDGMENTS

First and foremost, I would like to thank God for everything.

I wish to express my deepest gratitude to Dr. José L. Zayas-Castro for providing constant encouragement and support throughout this process, and my profound thanks to Dr. Peter J. Fabri for being my guide in the healthcare domain. Also, many thanks to Dr. Michael X. Weng for his time and insight in helping me to complete this work.

Special thanks to Dr. Anita Callahan for introducing me to this area of research, Dr. Geoffrey Okogbaa for his continued support, and Professor Dolores Gooding for her wisdom.

Finally, this research could not have been undertaken and successfully completed without the understanding and patience of my family and friends. They have my deepest appreciation.

TABLE OF CONTENTS

LIST OF TABLES iii

LIST OF FIGURES v

ABSTRACT........ vi

CHAPTER ONE INTRODUCTION........ 1
1.1 The Health Care Sector........ 1
1.2 The Health Care Supply Chain (HCSC)........ 1
1.3 Graduate Medical Education (GME) 2
1.3.1 GME Funding Mechanism........ 2
1.4 Research Objectives........ 3
1.5 Thesis Organization........ 3

CHAPTER TWO LITERATURE REVIEW 5
2.1 Introduction........ 5
2.2 Skill Mix 5
2.2.1 Health Care 6
2.3 Modes/Methodologies........ 8
2.3.1 Survey Design and Data Collection........ 8
2.3.2 Modeling........ 9
2.4 Problem Statement and Objectives 9
2.5 Research Scope 10

CHAPTER THREE METHODOLOGY 11
3.1 Introduction........ 11
3.2 The Rationale of the Survey Method 12
3.3 Survey Instrument Elements 14
3.3.1 Internal and External Validity........ 16
3.3.2 Data Analysis Tools........ 17
3.4 Construction of the Model........ 20
3.4.1 Introduction........ 20
3.4.2 LP Formulation........ 21
3.5 Validation of the Model........ 23

CHAPTER FOUR RESULTS AND DISCUSSION 24
4.1 Introduction 24
4.2 Analysis and Verification of the Research Proposition 24
4.2.1 Sample Size 26
4.2.2 Univariate Analyses – Internal Medicine 26
4.2.3 Multivariate Analyses – Internal Medicine 27
4.2.4 Qualitative Analysis 35
4.3 Other Residency Programs at USF College of Medicine 35
4.4 LP Model Results 36
4.5 Discussion 39

CHAPTER FIVE CONCLUSIONS AND FUTURE RESEARCH 43
5.1 Introduction 43
5.2 Summary of Results 43
5.3 Contributions of Research 44
5.4 Future Research 45

REFERENCES 47

APPENDICES 50
Appendix A Residency Program Director Survey Instrument 51
Appendix B Linear Programming Model and Output 57
Appendix C Data Analysis of Residency Programs at USF College of Medicine 59
Appendix D Definitions and E/M Codes for Tasks/Procedures 64

LIST OF TABLES

Table 2.1 Skill Mix – Determinants, Requirements and Possible Interventions 6

Table 3.1 Formal representation of Kruskal-Wallis Test..18

Table 4.1 Summary of Survey Results...25

Table 4.2 Descriptive Statistics: Proficiency 25

Table 4.3 Kruskal-Wallis Test: Proficiency versus Provider....................................26

Table 4.4 Kruskal-Wallis Test: Proficiency versus Task 27

Table 4.5 Kruskal-Wallis Test: Proficiency versus Dimension................................27

Table 4.6 ANOVA: Proficiency versus Provider, Task, Dimension 30

Table 4.7 General Linear Model: Proficiency versus Provider, Task, Dimension....31

Table 4.8 Nested ANOVA: Proficiency versus Provider (Dimension), Task 31

Table 4.9 Tukey Pairwise Comparison for Provider...32

Table 4.10 Tukey Pairwise Comparison for Task...33

Table 4.11 Tukey Pairwise Comparison for Dimension...34

Table 4.12 Regression Model for Proficiency 34

Table 4.13 General Linear Model: Proficiency versus Provider, Task, Dimension (Residency Programs-COMED) 36

Table 4.14 LP Model Inputs – Healthcare System 1 36

Table 4.15 LP Model Results (Mean Proficiency) – Healthcare System 1 37

Table 4.16 LP Model Results (Median and Mode Proficiency) – Healthcare System 1 38

Table 4.17 LP Model Inputs – Healthcare System 2 38

Table 4.18 LP Model Results (Mean Proficiency) – Healthcare System 2 39

Table 4.19 LP Model Results (Median and Mode Proficiency) – Healthcare System 2 39

Table 5.1 Summary of Research Proposition and Analysis 44

LIST OF FIGURES

Figure 1.1 Health Care Supply Chain 1

Figure 3.1 Reengineering GME Problem Description..12

Figure 3.2 Proposition and Data Analysis Strategy 14

Figure 3.3 General Factorial Design 19

Figure 3.4 Nested-Factorial Design..20

Figure 4.1 Histogram of Residuals 28

Figure 4.2 Normal Probability Plot of Residuals 29

Figure 4.3 Normal Probability Plot of Raw Data 29

RE-ENGINEERING GRADUATE MEDICAL EDUCATION: AN ANALYSIS OF THE CONTRIBUTION OF RESIDENTS TO TEACHING HOSPITALS UTILIZING A MODEL OF AN INTERNAL MEDICINE RESIDENCY PROGRAM

Ian M. Elius

ABSTRACT

According to the Institute of Medicine (IOM), the U.S. health care delivery system does not provide consistent, high-quality medical care to all people all the time. As a significant component of the health care delivery system, the state of Graduate Medical Education in the United States has prompted much analysis in recent years due to the general view that desired and actual outcomes are increasingly at variance with each other. One area of focus has been the implications of change for provider credentialing and funding of graduate medical education.

With this research we test the hypothesis that residents perform valuable work in the teaching hospitals where they undergo training, to inform the issue regarding provider credentialing for residents. We developed a framework to compare second-year residents (PGY2), physician assistants with one year of experience, and nurse practitioners with one year of experience to measurably address the interchangeability of providers. Data was collected by obtaining expert opinions on the proficiency of the three provider options (resident, physician assistant, nurse practitioner) in performing a set of tasks/procedures by surveying the program directors of Internal Medicine residency programs in the United States. The other residency programs at the University of South Florida's College of Medicine were also surveyed to obtain measurable performance on the service providers.

Statistical tools were used to analyze the survey responses, aggregate patient data and salary data for each provider. The data analysis and summary indicated that residents displayed higher levels of proficiency than physician assistants and nurse practitioners for the tasks investigated. The proficiency values were utilized as parameters to build a mathematical programming model with the objective of maximizing total proficiency by determining the optimal number of service providers. The model was developed and tested using data from two healthcare systems.

This research demonstrates that residents perform productive work at teaching hospitals, within the scope of the tasks and dimensions evaluated. Additionally, residents work capabilities were considered in the development of a model that can be scaled to investigate questions regarding skill mix.

CHAPTER ONE

INTRODUCTION

1.1 The Health Care Sector

According to the Institute of Medicine (IOM), the U.S. health care delivery system does not provide consistent, high-quality medical care to all people all the time [1]. National health expenditures in the United States reached $1.6 trillion in 2002, or 14.9 percent of the nation's gross domestic product (GDP), increasing 9.3% from the previous year [2].

1.2 The Health Care Supply Chain (HCSC)

The graduate medical education component of the health care supply chain includes hospitals, providers (physicians, physician assistants, nurse practitioners), Medicare, regulatory agencies, accreditation bodies as depicted by Figure 1.1. The IOM

Figure 1.1 Graduate Medical Education Component of Health Care Supply Chain

has made several recommendations to redesign the health care delivery system of the United States. One of these recommendations addresses the preparation of the workforce in health care by convening a summit of leaders in the health professions to develop strategies for: 1) restructuring clinical education, and 2) assessing the implications of change for provider credentialing, funding and sponsorship of education programs [3].

1.3 Graduate Medical Education (GME)

Graduate Medical Education (GME) refers to the period in a physician-in-training's education after graduation from medical school, and serves as the preparation for the independent practice of medicine. This period is typically of four years duration, and physicians-in-training are referred to as *residents*. Residency programs operate within each area of medical specialty at designated academic health centers or teaching hospitals [4].

The Accreditation Council for Graduate Medical Education (ACGME) is a private professional accreditation agency responsible for the accreditation of over 7,000 residency programs. ACGME stakeholders include residency programs, member organizations, patients, government agencies, and the general public.

ACGME provides its stakeholders with the assurance that residency programs are compliant with an approved set of educational standards. The standards are developed and programs reviewed by experts in each specialty who form the Residency Review Committees (RRC) [4].

1.3.1 GME Funding Mechanism

GME is currently funded from a number of different sources: Medicare, Medicaid, patient-care revenue, faculty practice plans, grants, endowments, Department of Defense (DoD), Department of Veterans Affairs[1] (VA). The largest contributor to the cost of educating residents is Medicare, which provides 74% of the funding. Reimbursement from Medicare is categorized into Direct Graduate Medical Education (DGME) and Indirect Medical Education (IME) payments [5].

DGME payments are used to support: 1) overhead expenses for GME, 2) salaries and fringe benefits for residents, 3) some compensation for teaching physician time, and 4) costs of the administrative staff for GME. Total DGME payments to academic health centers were $2.7 billion in 2000 [5].

IME payments are made to the academic health centers to reimburse additional costs for: 1) more complicated cases, 2) additional tests ordered by residents as part of the learning process, 3) uncompensated care, and 4) reduced patient-care productivity by teaching staff members. IME payments totaled $5.1 billion to academic health centers in 2000 [5].

1.4 Research Objectives

The goals of this work are two-fold: first, to discover if residents perform productive work by exploring the interchangeability of different types of service providers in an academic health center/teaching hospital; and second, to recommend an optimal skill mix for a residency program based on the results of a model constructed from the data collection and analysis.

1.5 Thesis Organization

The organization of this thesis is as follows: Chapter Two reviews the work of other authors in the area of skill mix and funding models; Chapter Three explains the

[1] Medicaid operates in forty-three states; VA and DoD operate their own programs distinct from the remainder of the U.S.

concepts of statistical analysis, linear programming and survey design and validity used in this research; Chapter Four provides data analysis results and discussion; Chapter Five describes the conclusions, contributions and possible future work.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

Graduate Medical Education (GME) in academic health centers (AHC) has been under increasing pressure from legislative and accreditation entities to improve the quality of educational outcomes while reducing costs [19]. A number of alternatives have been proposed including: 1) restructuring the financing mechanism for GME, and 2) reorganizing the roles of the health care team. These would be accomplished by: 1) changing the Medicare allocation from an entitlement to an appropriations process, and 2) allowing alternative health professionals to perform more advanced tasks to help achieve the work hours and workload goals for the resident [19, 20].

2.2 Skill Mix

The maximization of output of the industrial workplace is largely attributed to the theory of scientific management, proposed by Frederick W. Taylor in 1903. This theory espoused the decomposition of specific tasks and the interchangeability of individuals performing those tasks [18]. Initially, the tenets of scientific management were primarily applied to lower skilled workers in mass production facilities but they are increasingly being applied to professionals. This is due to a shift in the mode of professional practice from individual entrepreneurial activity to multi-professional corporate sites, which has occurred over time to take advantage of technological advances and central administrative capabilities. Contemporary professionals are primarily employees of large scale, hierarchical firms in the fields of accounting, architecture, law, and health care [18].

2.2.1 Health Care

An overview of skill mix in health care can be divided into three categories: 1) mix in nursing and other non-medical health professions, 2) introduction of new types of workers, and 3) role overlap between doctors and other health professionals [21]. Table 2.1 illustrates some determinants, requirements and possible interventions related to skill mix.

Table 2.1: Skill Mix – Determinants, Requirements and Possible Interventions [21]

Determinant	Requirement	Possible interventions
Cost containment	Improved management of organizational costs, specifically labor costs	Reduce unit labor costs or improve productivity by altering staff mix or level
Quality improvement	Improved quality of care	Improve use and deployment of staff skills to achieve best mix
New health sector programs or initiatives	Maximum health benefits of program implementation, by having appropriately skilled workers in place	Determine the cost-effective mix of staff required; enhance skills of current staff; introduce new types of worker
Health sector reform	Cost containment, improvements in quality of care and performance, and responsiveness of health sector organizations	Adjust staff roles; introduce new skills and new types of worker
Changes in the legislative/regulatory environment	Scope for changes in (or constraints on) role for different occupations, professions	Adjust staff roles; introduce new skills and new types of worker

In the first category, the most common model used is the qualified/unqualified mix with technical aides or vocationally trained assistants used to perform simple nursing tasks. The results in these cases are mixed, with a majority indicating positive outcomes with respect to cost containment. It should be noted that the majority of the literature in this area is published by supporters of health support staff [21]. In the second category, the most visible type of new worker has been the physician assistant whose role is uniquely to perform specific tasks previously performed solely by physicians. There is a high-growth trend in the occupation with 63,033 employed in 2002 and a projected 93,827 to be employed in 2012 [22]. In the third category, role overlap between doctors and nurses has been studied extensively with the resulting view that nurses should take on more advanced responsibilities. In many cases the outcomes were that nurses provided a more cost-effective alternative to doctors, and patient satisfaction reports were higher due to the extra time nurses spent with the patients [21].

More specific to skill mix with respect to residents, a survey conducted in 1995 by the Council of Teaching Hospitals indicated that 62% of its members employed physician assistants and assistant registered nurse practitioners to undertake tasks previously performed by residents [20]. Hospitals were seeking alternative health professionals to residents based on the projected shift in the nation's physician supply by the Council for Graduate Medical Education (COGME) at that time and the consequent reduction in the number of residents that would be available [20].

A meta-analysis of the literature conducted in 2003 sought to quantify and categorize the activities that residents engaged in at a teaching hospital. The findings indicated that residents time were allotted as follows: 36% on inpatient care, 15% on teaching and learning, 35% on tasks of marginal or no educational value, and 16% on 'other' or miscellaneous activities [20]. It should be noted at this point that grouping the 'inpatient care' and 'teaching and learning' categories together, and then grouping the 'marginal' and 'miscellaneous' categories will yield a 51%-51%[2] separation that supports

[2] Percentage exceeds 100% due to rounding error.

a major assumption of this research work of a 50%-50% split between work and study for the residents.

2.3 Modes/Methodologies

2.3.1 Survey Design and Data Collection

Survey instruments have long been used as data collection tools for various types of research studies, including those undertaken in the health care domain. In particular, the self-administered questionnaire has been used to obtain data from geographically diverse populations and also as a supplement to other data collection techniques. The major source of concern regarding survey instruments relate to instrument reliability and validity. Reliability refers to the consistency of the data obtained, and validity refers to the accuracy of the data [25].

A crucial aspect of the instrument design is the selection of a measurement scale to record the responses for each survey item. The visual analog scale (VAS) is a type of rating scale that allows the respondent to rank his/her preference and indicate this preference on a scale or line. The VAS derives its theoretical underpinnings from two sources: 1) the decision sciences/economics disciplines, which interpret the VAS scores as a measurable value function representing the strength of preferences under uncertainty distinct from a utility function, and 2) the psychology/psychophysics disciplines, which focus on the effect of stimuli and response modes on judgment [28]. A comparison between the VAS and the 7-point Likert scale showed that VAS had a larger response base than the Likert scale after standardizing on a 10-point scale. However, the variability was greater with VAS and the difference in responses was not statistically significant. Therefore, the two methods are comparable to represent results [27]. Another comparison including a VAS, a 12-point Borg scale, and a 5-point Likert scale concluded that the best overall scale was the VAS for reproducibility and sensitivity [26].

2.3.2 Modeling

The field of engineering has an opportunity to provide value to health care delivery by exploring the following capabilities: 1) developing better metrics, 2) identifying proven tools and methodologies for application in health care, and 3) developing quantitative models for more thorough examination and optimization of system performance [23].

One possible area of emphasis is the utilization of mathematical programming models applied to various aspects of health care. It is becoming increasingly important to efficiently allocate healthcare resources, and mathematical programming can be particularly useful in budget allocations, scheduling clinical studies and assigning medical personnel [15].

2.4 Problem Statement and Objectives

The literature addresses the need for re-engineering GME, and the complexities of the current GME funding mechanism [19]. Additionally, research efforts have been undertaken to provide the aggregate quantity and grouping of activities that residents perform in teaching hospitals [20]. It has also been demonstrated that alternative health professionals have been utilized to substitute for residents in performing tasks at a teaching hospital at a time when the physician workforce outlook was uncertain in the United States [21].

There appears to be a gap in the literature and corresponding knowledge on the topic of specific interchangeability among residents and alternative health professionals, which requires creating a framework to facilitate a standard measure for each type of service provider for specific tasks.

Survey methodology is utilized to obtain the data in this research effort, and the literature shows that these techniques are established in the health care domain as well as in other fields [24]. Additionally, linear programming is demonstrated as a tool with increased possibilities of application in the health care sector. Traditional areas include nurse scheduling, revenue management, hospital bed capacity modeling, reducing patient wait times, and optimizing inventory levels [15,16,17].

It appears from the literature that linear programming has not been used to model skill mix among providers in a teaching hospital. This research effort will utilize this modeling approach with its inherent flexibility to allow various scenarios to be easily evaluated while yielding optimal results, including scalability and transferability.

2.5 Research Scope

The research addresses the creation of a framework for data collection on three types of service providers, the analysis of survey responses to describe the differences among the providers, and the development of a residency program model to determine the optimal skill mix.

Survey methodology is used for the data collection phase of the research, and the findings and resulting conclusions are made with respect to residents and alternative health professionals in Internal Medicine residency programs.

CHAPTER THREE

METHODOLOGY

3.1 Introduction

Residents gain proficiency in advanced medical training through repetitive performance of routine tasks and exposure to complex cases. Other medical personnel would undertake these necessary tasks if residents were unavailable, distinguishing task completion from education. This leads to the notion that residents' responsibilities can be categorized as work or study.

We propose that under the 80-hour weekly limit rule instituted by ACGME, 40 hours is allotted to work and 40 hours is allotted to study. The productive work done by the residents may be compensated separately. As the residents progress through the residency program, they may achieve certification to perform various procedures and thus acquire the status of a professional.

In our study, we compare the proficiency of a second-year resident (PGY2), a Physician Assistant (PA) with one year of experience, and a Nurse Practitioner (NP) with one year of experience. The components of the research design are presented in Figure 3.1, beginning with the problem statement and ending with projected contributions.

Problem: Healthcare costs are spiralling upward in the United States. A portion of this cost is an entitlement paid by Medicare for the training of physicians in teaching hospitals. There is much debate about the financial model that should be used to support physicians-in-training. Medicare finances 75% of cost of GME for training purposes.

Sub-Problem: There is a lack of knowledge regarding the amount contributions that residents make to the teaching hospitals where they undergo training.

Hypothesis: Residents perform valuable work in the teaching hospitals where they undergo training. The training that residents undergo and the work performed are confounded.

Method to test hypothesis: Compare the interchangeability of health care providers (residents, PA's, NP's) for a finite number of tasks to quantify the work contributions of residents in a teaching hospital setting.

Assumptions: Work is necessary (ie. must be performed by some entity/resource/provider).
There is a 50%-50% split between valuable work and training/learning/study of the time residents spend at the teaching hospitals .
Residents (Second Yr), Physician Assistants (1 Yr Exp), Nurse Practitioners (1 Yr Exp) are comparable.

Instruments: Survey questionnaire.
Aggregate patient data categorized per task.
Salary data for each health care provider from a teaching hospital.

Tools: Statistical Analysis (Descriptive Statistics, Analysis of Variance (ANOVA)).
Mathematical Programming (LP).

Contributions: Demonstrate that residents perform productive work at teaching hospitals.
Provides a model that considers resident work capabilities to determine staffing needs in teaching hospitals.
Establishes a foundation for many questions regarding skill mix to be explored.

Figure 3.1: Re-engineering GME Research Design

The analytical framework utilized to undertake this research endeavor was the survey approach, combined with the quantitative tools that are used extensively in experimental design.

3.2 The Rationale of the Survey Method

The survey is a "system for collecting information from or about people to describe, compare, or explain their knowledge, attitudes or behavior". To establish the necessary rigor, seven components must be included in the survey method: setting objectives for information collection, survey design, preparing a valid and reliable instrument, administering the survey, survey data management and analysis, results reporting [24].

As outlined in Chapter One, the research questions relating to this work are: 1) to discover if residents perform productive work by exploring the interchangeability of different types of service providers in an academic health center/teaching hospital, and 2) to recommend an optimal skill mix for a residency program based on the results of a model constructed from data collection and analysis.

The proposition, primary data analysis strategy, and hypotheses are provided in Figure 3.2.

Proposition 1	The proficiency of residents, physician assistants and nurse practitioners are measurable and can be equated for specific procedures/tasks.	ANOVA	For 5 tasks measured in 3 dimensions: 15 hypotheses.

Hypothesis 1	For Task 1 along Dimension 1: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 2	For Task 1 along Dimension 2: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 3	For Task 1 along Dimension 3: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 4	For Task 2 along Dimension 1: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 5	For Task 2 along Dimension 2: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 6	For Task 2 along Dimension 3: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 7	For Task 3 along Dimension 1: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 8	For Task 3 along Dimension 2: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 9	For Task 3 along Dimension 3: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 10	For Task 4 along Dimension 1: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 11	For Task 4 along Dimension 2: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 12	For Task 4 along Dimension 3: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 13	For Task 5 along Dimension 1: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 14	For Task 5 along Dimension 2: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;
Hypothesis 15	For Task 5 along Dimension 3: Proficiency of residents, physician assistants and nurse practitioners are equal	ANOVA	$H_0: \mu_R = \mu_{PA} = \mu_{NP}$; $H_1: \mu_R ? \mu_{PA} ? \mu_{NP}$;

Figure 3.2: Proposition and Data Analysis Strategy

There are four types of survey instruments: self-administered questionnaires, interviews, structured record interviews, and structured observations [24]. Self-administered questionnaires were the type of survey instrument utilized in this research and a sample is shown in Appendix A.

3.3 Survey Instrument Elements

This research required expert opinions on the proficiency of the three provider options (resident, PA, NP) in performing a set of tasks/procedures. The survey instrument was developed, with proficiency measured on a visual analog scale from 0 – 100 with 100 representing the proficiency of an attending physician.

The tasks/procedures relate to five evaluation and management (E/M) codes[3] that are defined in Appendix D. The task/procedure titles follow:

- Inpatient Admission
- Inpatient Consultation
- Emergency Department Services
- Critical Care Evaluation
- Discharge Day Management [Hospital Discharge Services]

The proficiency of each task was measured along the dimensions of:

- Time (to completion)
- Quality (of outcome)
- Supervision (level required)
    - "Independence" was specifically measured in the questionnaire to maintain congruity in direction of increasing performance on the scale.

All Internal Medicine programs in the United States and all residency programs at the USF College of Medicine were surveyed. The survey instruments were distributed through the Graduate Medical Education (GME) Office at the USF College of Medicine.

[3] "The Evaluation and Management codes were first introduced in the 1992 CPT (Current Procedural Terminology), and were jointly developed by the AMA (American Medical Association) and HCFA (Health Care Financing Administration). The mutual goal was to provide Physicians and claim reviewers with advice on how to prepare or review documentation for evaluation and management services, and to increase accuracy and consistency in reporting levels of service furnished" [11].

Completed surveys were returned to the GME Office. Surveys were distributed on a second occasion to the Internal Medicine programs to improve participation.

Data related to the set of tasks/procedures indicated earlier was obtained from the USF Physicians Group for Tampa General Hospital and Moffitt Cancer Center by specifying the pertinent codes. Specifically, the number of tasks/procedures performed at each of these locations was provided on a quarterly basis for the one year. Also, salary information for each service provider type was provided.

3.3.1 Internal and External Validity

A descriptive factorial design was selected as the basis of the survey method. In this design the dependent variable was the program directors' expert opinions of provider proficiency. The independent variables or factors were: 1) provider (resident, physician assistant, or nurse practitioner), 2) task (inpatient admission, inpatient consultation, emergency department services, critical care evaluation, or discharge day management), 3) Dimension (time, quality, level of supervision).

This type of survey design has potential for invalidity in the following respects: internal (selection), and external (interactive effects of selection, reactive effects of testing, reactive effects of innovation) [25]. With regard to internal validity issues, selection of the survey participants (i.e. program directors of Internal Medicine residency programs) was based on the criteria that only they possessed the knowledge and organizational responsibility to rate the proficiency of each provider. Additionally, each program director had an equal, nonzero chance of participating because all program directors were surveyed. Regarding external validity issues: interactive effects of selection were not created due to the descriptive/non-experimental nature of the survey design; reactive effects of testing were not displayed because no pre-measures were done, therefore no portion of the participant population was sensitized; reactive effects of innovation were not created leading to uncharacteristic behavior due to the descriptive/non-experimental nature of the survey design.

3.3.2 Data Analysis Tools

Data analysis is defined as the examination, categorization, tabulation, or otherwise recombination of evidence to address the initial propositions of a study [9]. Non-parametric and parametric statistical tests were used as data analysis tools on the survey results. The proficiency values were treated as continuous variables, based on the use of a visual analog scale for collection purposes. The independent variables (service provider, task, and dimension) were treated as categorical variables, which require no special treatment to generalize known properties of regression to models of analysis of variance and covariance [12].

The Kruskal-Wallis test can be applied in the one factor ANOVA case. It is a non-parametric test for the situation where the ANOVA normality assumptions may not apply, and was used in this work to test the difference in means independent of the normality assumptions associated with the ANOVA tests [14].

Let n_i (i = 1, 2,…,k) represent the sample sizes for each of the samples (k groups) in the data. For the service provider and dimension factors, k = 915; for the task factor, k = 549. Next, rank the combined sample. Then compute R_i = the sum of the ranks for group i. Then the Kruskal-Wallis test statistic is:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1) \qquad (1)$$

This statistic, shown in equation 1, approximates a chi-square distribution with k-1 degrees of freedom if the null hypothesis of equal populations is true. The minimum sample size must be at least 5 for the approximation to be valid. We reject the null hypothesis of equal population means if the test statistic H is greater than $\chi^2_{\alpha, k-1}$ where χ^2 is the chi-square function.

A more formal description is shown in Table 3.1.

Table 3.1: Formal representation of Kruskal-Wallis Test

H_0:	$m_1 = m_2 = ... = m_k$
H_A:	$m_i \neq m_j$; for at least one set of i and j.
Test Statistic:	$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$
Significance Level:	a , typically set to 0.05.
Critical Region:	$H > \chi^2 a_{,k-1}$ where χ^2 is the chi-square function.
Conclusion:	Reject the null hypothesis if the test statistic lies in the critical region

Analysis of variance (ANOVA) is used to distinguish the main and interaction effects of categorical independent variables on an interval dependent variable [14]. The general factorial design and nested-factorial designs were used as data analysis tools as shown in Figures 3.3 and 3.4.

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \varepsilon_{ijkl}$$

where,

$i = 1, 2, \ldots, a$
$j = 1, 2, \ldots, b$
$k = 1, 2, \ldots, c$
$l = 1, 2, \ldots, n$

i = service providers; a=3;

1 = residents
2 = physician assistants
3 = nurse practitioners

j = tasks (procedures); b=5;

1 = Inpatient Admission
2 = Inpatient Consultation
3 = Emergency Department Services
4 = Critical Care Evaluation
5 = Discharge Day Management [Hospital Discharge Services]

k = dimensions; c=3;
1 = time
2 = quality
3 = level of supervision

l = replicates; n=61;

Figure 3.3: General Factorial Design

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_{k(i)} + (\tau\beta)_{ij} + (\beta\gamma)_{jk(i)} + \varepsilon_{ijkl}$$

where,

i = 1, 2,…, a
j = 1, 2,…, b
k = 1, 2,…, c
l = 1, 2,…, n

i = service providers; a=3;

1 = residents
2 = physician assistants
3 = nurse practitioners

j = tasks (procedures); b=5;

1 = Inpatient Admission
2 = Inpatient Consultation
3 = Emergency Department Services
4 = Critical Care Evaluation
5 = Discharge Day Management [Hospital Discharge Services]

k = dimensions; c=3;
1 = time
2 = quality
3 = level of supervision

l = replicates; n=61;

Figure 3.4: Nested-Factorial Design

3.4 Construction of the Model

3.4.1 Introduction

The results of the data analysis were utilized as parameters for the relevant variables necessary to build a mathematical programming model. The linear model served as a decision making tool to determine the optimal combination of service provider types in an Internal Medicine residency program. The service providers being evaluated were second year residents (R), physician assistants with one year of experience (PA), and nurse practitioners with one year of experience (NP). Physician assistants and nurse practitioners are independently certified professionals hired by hospitals (teaching and non-teaching) to perform specific tasks. Five of these tasks, which have previously been identified, were utilized in the data collection portion of this work. We make the assumption that service providers contribute an equal proportion of their time to tasks that are not included in the model.

3.4.2 LP Formulation

The objective of the model is to maximize proficiency, while meeting the cost and other requirements of the internal medicine residency program. Values of combined proficiency (P_C) for each type of service provider were extracted from the analysis of the data collected. The combined proficiency values are the mean of the proficiencies across three dimensions (time, quality, level of supervision) for each type of service provider. Consequently, the objective function of the LP model follows:

$$\text{Max } P_C = P_{C1}\, R + P_{C2}\, PA + P_{C3}\, NP \qquad (2)$$

The constraints of the model were constructed from various characteristics of a functioning Internal Medicine Residency Program at a healthcare system.

$$S_1\, R + S_2\, PA + S_3\, NP \leq \text{Budget per year} \qquad (3)$$

Equation 3 represents the financial constraints of the residency program. S_i, where i=1,…,3, represent the salaries of the service providers including benefits.

$$R \leq \text{Maximum \# of resident slots} \qquad (4)$$

$$R \leq 4\ AP \qquad (5)$$

Equation 4 and 5 represent the regulatory and accreditation constraints. Equation 4 denotes the number of resident slots allocated to a residency program that Medicare is willing to fund. Equation 5 characterizes the ratio of residents to attending physicians (AP) that are allowed by the accreditation body for residency programs (ACGME). There is no specific number indicated for a general Internal Medicine residency program. A majority of general residency programs have a 4:1, resident to attending physician ratio. Some specialties of Internal Medicine have a 1.5:1, resident to attending physician ratio.

$$U_1R + U_2PA + U_3NP \leq \text{max \# of tasks performed per year} \qquad (6)$$

Equation 6 represents the physical capacity constraints. U_i, where i=1,…,3, denotes the utility of the service provider that is computed by dividing the number of minutes worked in a year by the result of the division of the standard time an attending physician takes to perform a procedure/task (30 minutes) by the time proficiency of that provider.

An example will illustrate:

- Number of minutes worked in a year

  = (40 hrs/wk * 50 wks * 60 mins/hr) = 120, 000 minutes

- Proficiency of resident to attending physician = 75/100 = 0.75
- Standard time an attending physician takes to perform a procedure = 30 minutes
- U_1 = 120000 / (30 / 0.75) = 3000 procedures/yr

$$R \geq 0, PA \geq 0, NP \geq 0 \qquad (7)$$

Equation 7 represents the non-negativity constraints, so that a feasible solution would not be returned with a negative value for any type of service provider. Additionally, the variables for residents, physician assistants and nurse practitioners were set to return general integer values.

3.5 Validation of the Model

The model was validated by utilizing one set of data from a healthcare system in Tampa, Florida to build the model, then testing the model outcomes using a data set from another academic health center.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Introduction

The process of data collection from Program Directors of Internal Medicine in the United States, and all residency programs at the USF College of Medicine was completed over an eight (8) month period. During the data collection period, the Internal Medicine Program Directors were surveyed on two occasions and all other residency programs at USF were surveyed once. The second survey to the Internal Medicine residency programs was sent to the non-respondent Program Directors, in order to generate increased participation. The survey instrument was sent to 389 Internal Medicine Programs, 98 (25.2%) total responses were received with 61 (15.7%) responses completed. At the USF College of Medicine, the survey instrument was sent to 45 residency programs (including specialties, and sub-specialties) with 11 (24.4%) total responses received and 5 (11.1%) responses completed. These responses were analyzed using statistical tests, specifically the Kruskal-Wallis and Analysis of Variance (ANOVA) tests to support the research proposition.

4.2 Analysis and Verification of the Research Proposition

The research proposition statement follows:

The proficiency of residents, physician assistants, and nurse practitioners are measurable and can be equated for specific tasks/procedures.

To support this proposition a total of fifteen (15) hypotheses were made: one hypothesis for each of the five (5) tasks across the three (3) dimensions. For example, the hypothesis for Inpatient Admission is that each service provider has the same level of proficiency in terms of time, quality, and level of supervision required. Table 4.1 shows the summary of survey results with the mean, median and mode values for each type of service provider per task and dimension.

Table 4. 1 Summary of Survey Results

Survey Question	mean			median			mode		
	R	PA	NP	R	PA	NP	R	PA	NP
Time Task 1	71	58	56	75	62	56	75	63	50
Time Task 2	64	49	47	65	50	50	75	50	50
Time Task 3	69	58	55	73	63	53	75	75	50
Time Task 4	68	48	45	69	50	47	50	50	50
Time Task 5	73	67	68	75	71	71	75	75	75
Quality Task 1	79	61	59	76	63	56	75	50	50
Quality Task 2	74	52	51	75	52	50	75	50	50
Quality Task 3	74	59	56	75	63	59	75	75	50
Quality Task 4	76	50	48	75	50	50	75	50	50
Quality Task 5	77	70	72	76	75	75	75	75	75
Supervision Task 1	76	55	53	75	56	54	75	50	50
Supervision Task 2	71	48	46	75	50	50	75	53	50
Supervision Task 3	72	54	51	75	54	50	75	50	50
Supervision Task 4	70	45	42	75	47	47	75	25	25
Supervision Task 5	76	69	69	75	74	74	75	75	75
Grand Mean/Median/Mode	73	56	55	75	58	54	75	50	50

Table 4.2 shows the descriptive statistics for the proficiency variable in the complete data set for the Internal Medicine residency programs.

Table 4.2 Descriptive Statistics: Proficiency

```
Variable            N       Mean     Median     TrMean      StDev    SE Mean
Proficiency      2745     61.185     64.000     62.234     21.380      0.408

Variable      Minimum    Maximum         Q1         Q3
Proficiency     0.000    100.000     50.000     75.000
```

4.2.1 Sample Size

With a response rate of 61 from a population of 389 Internal Medicine Program Directors surveyed, the confidence interval was computed as ±11.54 (~±12) with a confidence level of 95%. For the USF College of Medicine residency program responses of 5 from a population of 45, the confidence interval was computed as ±41.79 (~±42) with a confidence level of 95%.

4.2.2 Univariate Analyses – Internal Medicine

The Kruskal-Wallis test was used to evaluate the level of significance for proficiency in terms of service provider type, task, and dimension respectively. Tables 4.3, 4.4 and 4.5 show the results of each test.

Table 4.3 Kruskal-Wallis Test: Proficiency versus Provider

```
Kruskal-Wallis Test on Proficiency

Provider    N    Median    Ave Rank         Z
1         915     75.00      1812.2     20.53
2         915     58.00      1184.5     -8.81
3         915     54.00      1122.3    -11.72
Overall  2745                1373.0

H = 424.34  DF = 2  P = 0.000
H = 425.58  DF = 2  P = 0.000 (adjusted for ties)
```

The service providers are statistically significantly different from each other with a p-value of 0.000, with residents (provider 1) having the highest ranking followed by physician assistants and nurse practitioners in descending order.

Table 4.4 Kruskal-Wallis Test: Proficiency versus Task

Kruskal-Wallis Test on Proficiency

Task	N	Median	Ave Rank	Z
1	549	65.00	1421.6	1.61
2	549	56.00	1160.6	-7.02
3	549	63.00	1349.1	-0.79
4	549	54.00	1152.6	-7.28
5	549	75.00	1781.1	13.49
Overall	2745		1373.0	

H = 230.01 DF = 4 P = 0.000
H = 230.69 DF = 4 P = 0.000 (adjusted for ties)

The tasks are statistically significantly different from each other with a p-value of 0.000, with Discharge Day Management [Hospital Discharge Services] (task 5) having the highest ranking followed by Inpatient Admission, Emergency Department Services, Inpatient Consultation and Critical Care Evaluation in descending order.

Table 4.5 Kruskal-Wallis Test: Proficiency versus Dimension

Kruskal-Wallis Test on Proficiency

Dimension	N	Median	Ave Rank	Z
1	915	63.00	1283.2	-4.20
2	915	67.00	1474.2	4.73
3	915	64.00	1361.6	-0.53
Overall	2745		1373.0	

H = 26.85 DF = 2 P = 0.000
H = 26.93 DF = 2 P = 0.000 (adjusted for ties)

The dimensions are statistically significantly different from each other with a p-value of 0.000, with quality (dimension 2) having the highest ranking followed by level of supervision and time in descending order.

4.2.3 Multivariate Analyses – Internal Medicine

The Analysis of Variance (ANOVA) tests were performed to analyze the simultaneous effect of all of the factors on proficiency. The validity of the assumptions required for the ANOVA to be an exact test of the hypothesis of no difference in means was checked. These assumptions are: 1) the observations are adequately described by the fixed effects model in Figures 3.3 and 3.4, and 2) the errors are normally independently distributed with mean zero and constant but unknown variance σ^2 [NID $(0, \sigma^2)$]. The normality assumption was checked by plotting a histogram of the residuals, which resulted in a plot similar to a sample from a normal distribution centered at zero indicating normality as shown in Figure 4.1. Additionally, as depicted in Figures 4.2 and 4.3 normal probability plots of the residuals and the raw data were constructed with straight lines approximated in both plots indicating normality.

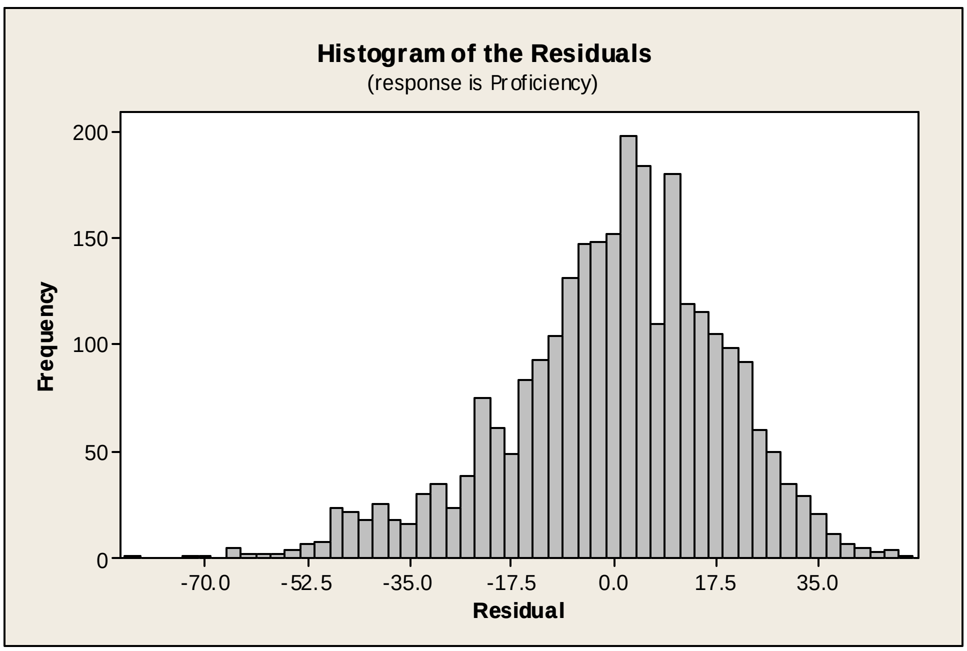


Figure 4.1 Histogram of Residuals

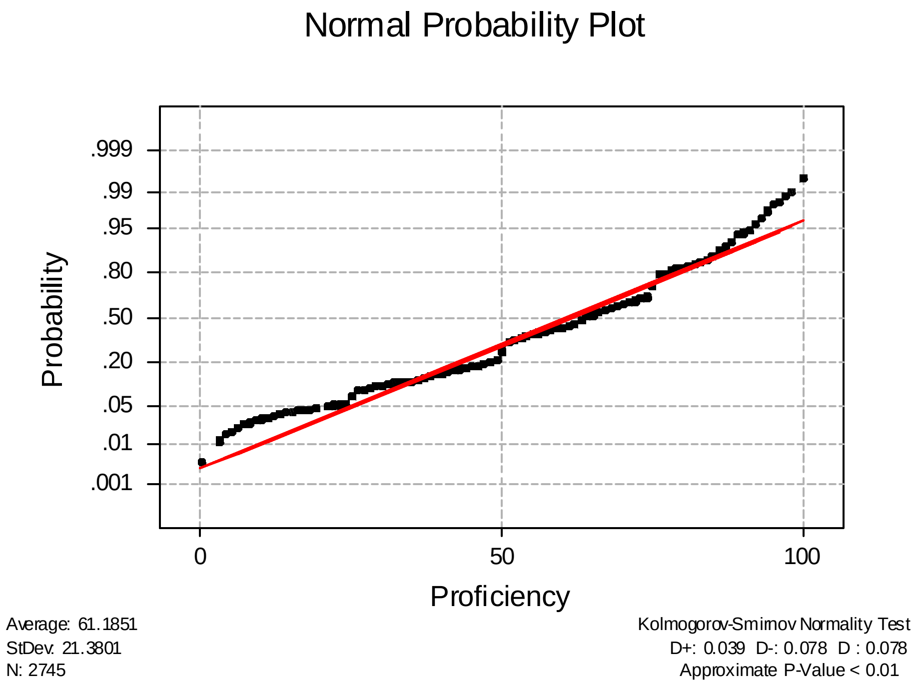


Figure 4.2 Normal Probability Plot for Residuals

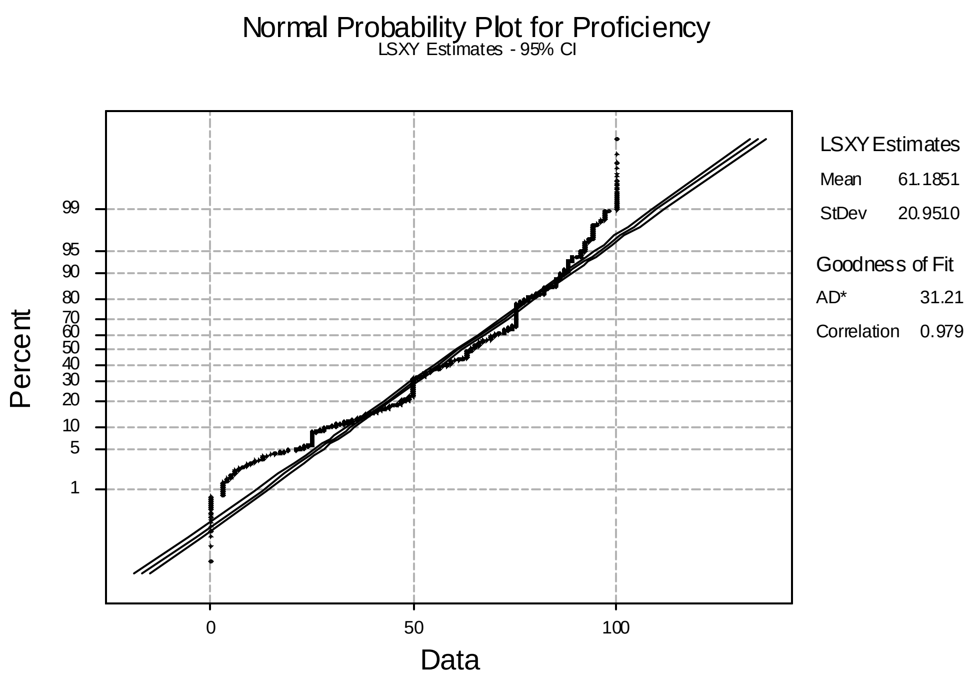


Figure 4.3 Normal Probability Plot of Raw Data

The full factorial ANOVA in Table 4.6 shows that the main factors provider, task, and dimension are statistically significant, with test statistics for each of the factors supporting the results of the univariate analyses.

Table 4.6 ANOVA: Proficiency versus Provider, Task, Dimension

Factor	Type	Levels	Values				
Provider	fixed	3	1	2	3		
Task	fixed	5	1	2	3	4	5
Dimension	fixed	3	1	2	3		

Analysis of Variance for Proficiency

Source	DF	SS	MS	F	P
Provider	2	181686	90843	257.12	0.000
Task	4	94871	23718	67.13	0.000
Dimension	2	11070	5535	15.67	0.000
Error	2736	966675	353		
Total	2744	1254302			

MANOVA for Provider   s = 1   m = 0.0   n = 1367.0

Criterion	Test Statistic	F	DF	P
Wilk's	0.84179	257.115	(2, 2736)	0.000
Lawley-Hotelling	0.18795	257.115	(2, 2736)	0.000
Pillai's	0.15821	257.115	(2, 2736)	0.000
Roy's	0.18795			

MANOVA for Task   s = 1   m = 1.0   n = 1367.0

Criterion	Test Statistic	F	DF	P
Wilk's	0.91063	67.129	(4, 2736)	0.000
Lawley-Hotelling	0.09814	67.129	(4, 2736)	0.000
Pillai's	0.08937	67.129	(4, 2736)	0.000
Roy's	0.09814			

MANOVA for Dimension   s = 1   m = 0.0   n = 1367.0

Criterion	Test Statistic	F	DF	P
Wilk's	0.98868	15.667	(2, 2736)	0.000
Lawley-Hotelling	0.01145	15.667	(2, 2736)	0.000
Pillai's	0.01132	15.667	(2, 2736)	0.000
Roy's	0.01145			

The full factorial ANOVA in Table 4.7 shows that the main factors (provider, task, and dimension) are statistically significant. Also, the provider/task and provider/dimension interactions are significant. The task/dimension interaction and provider/task/dimension interactions are not significant. Given these results, a nested-factorial ANOVA was computed in Table 4.8 with dimension nested within provider. Using this statistical model, all factors are significant.

Table 4.7 General Linear Model: Proficiency versus Provider, Task, Dimension

Factor	Type	Levels	Values
Provider	fixed	3	1 2 3
Task	fixed	5	1 2 3 4 5
Dimension	fixed	3	1 2 3

Analysis of Variance for Proficiency, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Provider	2	181686.1	181686.1	90843.1	262.17	0.000
Task	4	94870.7	94870.7	23717.7	68.45	0.000
Dimension	2	11070.5	11070.5	5535.2	15.97	0.000
Provider*Task	8	24755.7	24755.7	3094.5	8.93	0.000
Provider*Dimension	4	3837.8	3837.8	959.5	2.77	0.026
Task*Dimensio	8	1643.5	1643.5	205.4	0.59	0.785
Provider*Task*Dimension	16	889.2	889.2	55.6	0.16	1.000
Error	2700	935548.5	935548.5	346.5		
Total	2744	1254302.0				

Table 4.8 Nested ANOVA: Proficiency versus Provider (Dimension), Task

Factor	Type	Levels	Values
Provider	fixed	3	1 2 3
Dimension(Provider)	fixed	9	1 2 3 1 2 3 1 2 3
Task	fixed	5	1 2 3 4 5

Analysis of Variance for Proficiency, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Provider	2	181686	181686	90843	257.76	0.000
Dimension(Provider)	6	14908	14908	2485	7.05	0.000
Task	4	94871	94871	23718	67.30	0.000
Error	2732	962837	962837	352		
Total	2744	1254302				

Tukey’s test was used for all pairwise mean comparisons showing that residents have a higher proficiency than physician assistants and nurse practitioners, as shown in Table 4.9. Additionally, physician assistants and nurse practitioners are not significantly different from each other in terms of proficiency.

Table 4.9 Tukey Pairwise Comparison for Provider

```
Tukey 95.0% Simultaneous Confidence Intervals
Response Variable Proficiency
All Pairwise Comparisons among Levels of Provider

Provider = 1 subtracted from:

Provider     Lower    Center     Upper  ----+---------+---------+---------+--
2           -18.46    -16.40    -14.35     (---*--)
3           -20.06    -18.00    -15.95   (--*--)
                                        ----+---------+---------+---------+--
                                          -18.0     -12.0      -6.0       0.0


Provider = 1 subtracted from:

Level        Difference       SE of             Adjusted
Provider       of Means  Difference   T-Value    P-Value
2                -16.40      0.8777    -18.69     0.0000
3                -18.00      0.8777    -20.51     0.0000

Provider = 2 subtracted from:

Level        Difference       SE of             Adjusted
Provider       of Means  Difference   T-Value    P-Value
3                -1.597      0.8777    -1.819     0.1633
```

The Tukey pairwise comparison test was also used for analyzing tasks and dimensions. As shown in Table 4.10, Discharge Day Management (task 5) had the highest proficiency, followed by Inpatient Admission (task 1) and Emergency Department Services (task 3), then followed by Inpatient Consultation (task 2) and Critical Care Services (task 4). Table 4.11 shows that quality (dimension 2) had the highest overall proficiency, followed by time (dimension 1) and level of supervision (dimension 3).

Table 4.10 Tukey Pairwise Comparison for Task

```
Tukey 95.0% Simultaneous Confidence Intervals
Response Variable Proficiency
All Pairwise Comparisons among Levels of Task

Task = 1 subtracted from:

Task     Lower    Center     Upper  --+---------+---------+---------+----
2       -10.17    -7.078    -3.982    (--*--)
3        -5.48    -2.384     0.712         (--*--)
4       -11.56    -8.466    -5.370  (---*--)
5         4.96     8.053    11.149                   (--*--)
                                    --+---------+---------+---------+----
                                     -10         0        10        20


Task = 1 subtracted from:

Level    Difference       SE of             Adjusted
Task       of Means  Difference   T-Value    P-Value
2            -7.078       1.135    -6.239     0.0000
3            -2.384       1.135    -2.102     0.2193
4            -8.466       1.135    -7.462     0.0000
5             8.053       1.135     7.098     0.0000

Task = 2 subtracted from:

Level    Difference       SE of             Adjusted
Task       of Means  Difference   T-Value    P-Value
3             4.694       1.135     4.137     0.0003
4            -1.388       1.135    -1.223     0.7377
5            15.131       1.135    13.337     0.0000

Task = 3 subtracted from:

Level    Difference       SE of             Adjusted
Task       of Means  Difference   T-Value    P-Value
4            -6.082       1.135    -5.361     0.0000
5            10.437       1.135     9.200     0.0000

Task = 4 subtracted from:

Level    Difference       SE of             Adjusted
Task       of Means  Difference   T-Value    P-Value
5             16.52       1.135     14.56     0.0000
```

Table 4.11 Tukey Pairwise Comparison for Dimension

```
Tukey 95.0% Simultaneous Confidence Intervals
Response Variable Proficiency
All Pairwise Comparisons among Levels of Dimension

Dimension = 1 subtracted from:

Dimension    Lower    Center     Upper  --------+---------+---------+--------
2            2.296    4.3530     6.410                         (----*-----)
3           -1.864    0.1923     2.249              (-----*----)
                                        --------+---------+---------+--------
                                              -3.5       0.0       3.5

Dimension = 1 subtracted from:

Level        Difference       SE of             Adjusted
Dimension      of Means  Difference   T-Value    P-Value
2                4.3530      0.8788    4.9534     0.0000
3                0.1923      0.8788    0.2189     0.9739

Dimension = 2 subtracted from:

Level        Difference       SE of             Adjusted
Dimension      of Means  Difference   T-Value    P-Value
3                -4.161      0.8788    -4.735     0.0000
```

Regression analysis was used to develop an empirical model for proficiency as depicted in Table 4.12, however the predictive value of the model is severely limited due to the qualitative nature of the main factors.

Table 4.12 Regression Model for Proficiency

```
The regression equation is
Proficiency = 74.6 - 9.00 Provider + 1.47 Task + 0.096 Dimension

Predictor        Coef     SE Coef         T        P       VIF
Constant       74.579       1.595     46.75    0.000
Provider      -9.0005      0.4670    -19.27    0.000       1.0
Task           1.4718      0.2696      5.46    0.000       1.0
Dimensio       0.0962      0.4670      0.21    0.837       1.0

S = 19.98              R-Sq = 12.8%         R-Sq(adj) = 12.7%
PRESS = 1097253        R-Sq(pred) = 12.52%

Analysis of Variance

Source            DF          SS          MS         F         P
Regression         3      160157       53386    133.74     0.000
Residual Error  2741     1094145         399
Total           2744     1254302
```

4.2.4 Qualitative Analysis

The survey instrument contained sixteen (16) questions, with the last question being open-ended in nature to elicit descriptive responses from the participants. The participants were asked to suggest any other possible factors to evaluate proficiency of the service providers in addition to time, quality and level of supervision. There were a range of responses, however, the highest frequency responses related to patient satisfaction and professionalism.

4.3 Other Residency Programs at USF College of Medicine

The data obtained from the survey responses at the USF College of Medicine residency programs were analyzed with the same tools utilized on the Internal Medicine residency program survey data. Assumptions of normality were tested; then the Kruskal-Wallis and ANOVA methods were used to test for differences in means among the factors; followed by the Tukey comparison test to specifically determine which factors differ significantly from each other. The full results of the analyses are shown in Appendix C.

The univariate analyses (Kruskal-Wallis tests) for proficiency versus each factor (provider, task, dimension) each showed a statistical significant difference from each other. The multivariate analyses supported these differences, as shown by the ANOVA test results in Table 4.13.

Table 4.13 General Linear Model: Proficiency versus Provider, Task, Dimension (Residency Programs-COMED)

```
Factor     Type Levels Values
Provider  fixed      3 1 2 3
Task      fixed      5 1 2 3 4 5
Dimension fixed      3 1 2 3

Analysis of Variance for Proficiency, using Adjusted SS for Tests

Source                      DF     Seq SS     Adj SS     Adj MS       F      P
Provider                     2     4641.3     4641.3     2320.7    6.37  0.002
Task                         4     3756.1     3756.1      939.0    2.58  0.039
Dimension                    2     2378.8     2378.8     1189.4    3.27  0.040
Provider*Task                8     3447.2     3447.2      430.9    1.18  0.312
Provider*Dimension           4      276.2      276.2       69.1    0.19  0.944
Task*Dimension               8     1748.3     1748.3      218.5    0.60  0.777
Provider*Task*Dimension     16      898.8      898.8       56.2    0.15  1.000
Error                      180    65560.8    65560.8      364.2
Total                      224    82707.6
```

The ANOVA analysis indicates that the main factors (provider, task, dimension) are significantly different, however, the interaction terms are not significantly different.

4.4 LP Model Results

The LP model generated optimal values for the number of service provider types given the constraints of the internal medicine program. The inputs for the model of healthcare system 1 are shown in Table 4.14, and the model results are displayed in Table 4.15. Residents are consistently recommended in greater numbers than physician assistants and nurse practitioners for an optimal skill mix. The output from the Lindo® linear program is shown in Appendix B.

Table 4.14 LP Model Inputs – Healthcare System 1

	Resident	Physician Assistant	Nurse Practitioner
Salaries + Benefits[4]/yr	$50,165	$83,820	$82,550
# Tasks /provider/yr	2,840	2,320	2,240
Medicare Slots	10		
Attending Physicians	5		
Provider Budget/yr	$14,200,000		
Max # Tasks /yr	24,461		

[4] Indicates average provider salaries for specific healthcare system for year 2004, plus 27% benefits.

Table 4.15 LP Model Results (Mean Proficiency) – Healthcare System 1

Service Provider	Resident	Physician Assistant	Nurse Practitioner	Combined Proficiency	Budget ($)	Tasks (#)
Proficiency	73	56	55	-	-	-
Model Skill Mix	7	1	1	622	517,525	24,440
Current Skill Mix	10	0	1	785	584,200	30,640

The model generates the combination of providers capable of performing the tasks while maximizing combined proficiency. The binding constraint for the model is the maximum number of tasks to be performed annually. In other words, the number of tasks is the only fact that affects the level of skill mix, and the current budget is more than what is needed. Therefore, the 'current skill mix' row of Table 4.15 shows a higher combined proficiency, budget allocation, and number of tasks performed than the model results. It is important to note at this point that the model only considers the optimal skill mix for the five tasks selected. The service providers currently in the residency program, in excess of the number generated by the model, are performing other tasks not considered in the model or pursuing educational objectives (in the case of residents). The binding constraint (maximum number of tasks) was relaxed to consider the combined proficiency and budget allocation values. The upper limit of the binding constraint is determined by facility and safety issues, which were outside the scope of this research.

What-if analyses were also performed to explore changes to the optimal solution given changes to parameters in the LP. The objective function coefficients were changed from mean values of proficiency to median and mode proficiency values. The summarized results are shown in Table 4.16. The combined proficiency values for both alternative scenarios are higher than the result obtained when using the mean proficiency values for each provider, primarily due to slightly higher values of proficiency for the resident in each case. The maximum number of tasks is the binding constraint in each case; however, the budget allocation is lower for the scenario with mode values because a physician assistant is replaced with a nurse practitioner who has slightly lower compensation.

Table 4.16 LP Model Results (Median and Mode Proficiency) – Healthcare System 1

Service Provider	Resident	Physician Assistant	Nurse Practitioner	Combined Proficiency	Budget ($)	Tasks (#)
Proficiency (Median)	75	58	54	-	-	-
Model Skill Mix	7	1	1	637	517,525	24,440
Proficiency (Mode)	75	50	50	-	-	-
Model Skill Mix	7	0	2	625	516,255	24,360

The model was tested by utilizing the characteristics from another healthcare system. The new parameters for the model inputs are shown in Table 4.17, and the results for Healthcare System 2 are shown in Table 4.18. The corresponding what-if analyses are summarized in Table 4.19. Healthcare System 2 has a smaller residency program than Healthcare System 1, therefore the number of resident slots, attending physicians, provider budget, and maximum number of tasks are reduced in magnitude.

Table 4.17 LP Model Inputs – Healthcare System 2

	Resident	Physician Assistant	Nurse Practitioner
Salaries + Benefits/yr	$50,165	$83,820	$82,550
# Tasks /provider/yr	2,840	2,320	2,240
Medicare Slots	4		
Attending Physicians	2		
Provider Budget/yr	$5,600,000		
Max # Tasks /yr	5,322		

The results generated by the model are shown in Table 4.18 in comparison with the current skill mix scenario. Two providers are reduced from the total number currently contributing to the Internal Medicine residency program. The maximum number of tasks continues to be the binding constraint, but is relaxed to determine the combined proficiency and budget allocation required for the current skill mix.

Table 4.18 LP Model Results (Mean Proficiency) – Healthcare System 2

Service Provider	Resident	Physician Assistant	Nurse Practitioner	Combined Proficiency	Budget ($)	Tasks (#)
Proficiency	73	56	55	-	-	-
Model Skill Mix	1	1	0	129	133,985	5,160
Current Skill Mix	4	0	0	292	200,660	11,360

In Table 4.19, the alternative scenarios used for Healthcare System 2 resulted in identical budget allocations and number of tasks completed because the skill mix generated by the model remained the same in both cases. The combined proficiency for the scenario with median values is higher than for the original model results using mean values because of the higher proficiency values for residents and physician assistants. The higher proficiency value for residents in the scenario with mode values is not sufficient to offset the lower proficiency of physician assistants in that scenario.

Table 4.19 LP Model Results (Median and Mode Proficiency) – Healthcare System 2

Service Provider	Resident	Physician Assistant	Nurse Practitioner	Combined Proficiency	Budget ($)	Tasks (#)
Proficiency (Median)	75	58	54	-	-	-
Model Skill Mix	1	1	0	133	133,985	5,160
Proficiency (Mode)	75	50	50	-	-	-
Model Skill Mix	1	1	0	125	133,985	5,160

4.5 Discussion

The findings of the analyses performed on the data collected through the surveys, and the optimal skill mix recommended by the model support the research proposition.

The univariate and multivariate statistical tests show a significant difference among the types of service providers, with pairwise comparisons indicating that residents have a higher proficiency than physician assistants and nurse practitioners (with a 95% confidence level). Also, physician assistants and nurse practitioners are not significantly different from each other indicating their interchangeability for the tasks investigated with no difference in proficiency levels.

The full factorial ANOVA showed significant interaction between provider and task, and provider and dimension. The provider and task interaction indicates that the effect of the provider on proficiency is dependent on the task being performed. Additionally, the effect of the provider on proficiency is dependent on the dimension being evaluated. The latter interaction is further supported statistically by the nested-factorial analysis with dimension nested within provider and task as main factors. There is no significant interaction between task and dimension, indicating independence between these two factors.

The results of the LP model are consistent with the factorial analyses, as residents are generated in greater numbers than the other service providers with the objective of maximizing proficiency. Additionally, it should be noted that residents have the lowest unit cost among the service providers therefore the model is validated logically by selecting the highest proficiency and lowest cost provider within the constraints of the program.

The model provides value to decision makers in hospitals by generating skill mix options based on the combined proficiency values of service providers. The binding constraint is the number of tasks to be performed; however, the flexibility of the model allows this constraint to be relaxed to explore scenarios unique to each residency program.

In Healthcare System 1, it is observed that a physician assistant was added to the skill mix, but that the total number of providers decreased by 2. The what-if scenarios generated the same total number of providers as the primary model, however, a nurse practitioner replaced a physician assistant when the proficiencies were the same. This result may be due to the model selecting the lower-cost option. In Healthcare System 2, a physician assistant was also added to the skill mix, but the total number of providers decreased by 2. The what-if scenarios provided the same results as the primary model for this Healthcare System.

It is important to reiterate that only the five tasks investigated in this research are considered in the model to determine the optimal number of service providers. The seemingly excess numbers of service providers currently in the residency programs perform tasks that are not included in this work. Therefore, the assumption that each service provider contributes the same proportion of his/her time to tasks not included in the model is necessary for consistency.

With the research proposition supported by the above data analyses and model, it can be implied that residents, physician assistants and nurse practitioners are interchangeable specifically for the five tasks selected in an internal medicine program. This measurable interchangeability enables hospital administrators to view residents as resources comparable to the other service providers, and to develop models to commercially justify their presence in a teaching hospital beyond the educational objectives of the residency program.

The analysis of the data from the other residency programs at the USF College of Medicine yielded similar results to the results from the internal medicine residency programs. There was a significant difference between providers, with residents demonstrating higher proficiency than physician assistants and nurse practitioners using pairwise comparison. Additionally, there were no significant interactions between provider and task and provider and dimension. These results are promising for the notion of generalizing the framework to model all residency programs.

The next chapter summarizes the research contributions, and outlines the possible future research directions from the foundation established by this work.

CHAPTER FIVE

CONCLUSIONS AND FUTURE RESEARCH

5.1 Introduction

Health care delivery in the United States has become an issue of the highest national priority, and strategies for reform are being sought in different components of the health care system. The issue of provider credentialing is a critical element in the effort to re-engineer graduate medical education, which will potentially have a significant impact on the health care system. This research was undertaken to analyze the contributions of residents to a teaching hospital in terms of work, and to provide the foundation for a new framework of provider credentialing and funding mechanism for GME.

Specifically, the goals of this work were two-fold: first, to discover if residents perform productive work by exploring the interchangeability of different types of service providers in an academic health center/teaching hospital; and second, to recommend an optimal skill mix for a residency program based on the results of a model constructed from the data collection and analysis.

5.2 Summary of Results

The primary conclusion of this research endeavor is that residents do perform work as measured by the set of tasks utilized and the dimensions evaluated, in comparison with physician assistants and nurse practitioners. Table 5.1 shows the results of the research proposition and hypothesis tests.

Table 5.1 Summary of Research Proposition and Analysis

Proposition	*The proficiency of residents, physician assistants, and nurse practitioners are measurable and can be equated for specific tasks/procedures.*	Supported
Factor	Type	Proficiency(Rank)
Provider	Resident	73 (1[5])
	Physician Assistant	56 (2)
	Nurse Practitioner	55 (2)
Task	Inpatient Admission	70 (2)
	Inpatient Consultation	63 (3)
	Emergency Department Services	67 (2)
	Critical Care Evaluation	63 (3)
	Discharge Day Management	73 (1)
Dimension	Time	65 (2)
	Quality	70 (1)
	Level of Supervision	67 (2)

5.3 Contributions of Research

This research contributes to re-engineering graduate medical education by providing a framework of analysis to quantify how well work is performed by service providers, by demonstrating that residents perform productive work at teaching hospitals, and by developing a model of an internal medicine program that generates the optimal

[5] Values in parentheses indicate relative ranking of factors. Factors of the same rank are not significantly different from each other.

skill mix of service providers. The measurable comparison of residents to commercial providers have implications for the manner in which residents are viewed by teaching hospitals, and creates a framework for provider credentialing and re-evaluating the GME funding mechanism.

The use of linear programming as a tool to model residency programs for skill mix optimality is novel. Linear programming has previously been utilized in the health care sector in the more traditional areas of staff scheduling and inventory management. LPs provide a level of flexibility for adapting to changing parameters among residency programs.

5.4 Future Research

There are directions to explore in pursuing extensions and refinements to this research endeavor. First, the framework of analysis for quantifying work performed by residents can be expanded. The tasks/procedures investigated may be expanded to be more relevant to the specific residency program being explored. Also, the initial additions to the dimensions of evaluation can be obtained from the most frequent responses to the last question on the survey instrument: “patient satisfaction” and “professionalism”. Second, proficiency values for each type of service provider, task and dimension can be obtained by survey methodology and/or other data collection techniques for other residency programs. Third, the LP model of the internal medicine residency programs can be applied to other residency programs.

This research effort took a critical first step in redefining residents’ role in a teaching hospital by measuring the work being performed. The efforts to extend this research in different directions will increase the likelihood of re-engineering graduate medical education in a manner that will positively impact the health care system.

REFERENCES

[1] Institute of Medicine. *Informing the Future: Critical Issues in Health. 2nd ed*. Washington, DC: The National Academies Press, 2003.

[2] *"Highlights - National Health Expenditures, 2002."* 17 Sep 2004. Centers for Medicare and Medicaid Services. U.S. Department of Health and Human Services. 17 Oct 2004 <http://www.cms.hhs.gov/statistics/nhe/historical/highlights.asp>.

[3] Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academy Press, 2001.

[4] "Accreditation Council for Graduate Medical Education (ACGME)." ACGME: Accreditation Council for Graduate Medical Education. Accreditation Council for Graduate Medical Education (ACGME). 19 Oct 2004 <http://www.acgme.org/acWebsite/about/ab_roleACGME.asp

[5] Covey, MD, MBA, Aaron S., and Gary E. Friedlaender, MD. "Financing Graduate Medical Education: Sorting out the confusion." *The Journal of Bone and Joint Surgery, Incorporated* 85-A.8 (2003): 1594-1604.

[6] Burns, Lawton R. . *The Health Care Value Chain: Producers, Purchasers, and Providers*. San Francisco: Wiley, 2002.

[7] Bromley, D. B. (1990). Academic contributions to psychological counselling: I. A philosophy of science for the study of individual cases. *Counselling Psychology Quarterly*, 3(3), 299-307

[8] Zucker, D. M. (2001, June). Using case study methodology in nursing research. [45 paragraphs]. *The Qualitative Report* [On-line serial], *6*(2). Available: http://www.nova.edu/ssss/QR/QR6-2/zucker.html

[9] Yin, R. K. (1994). Case study research: Design and methods (2nd ed.). Newbury Park, CA: Sage Publications.

[10] Tellis. W. (1997, September). Application of a case study methodology [81 paragraphs]. The Qualitative Report [On-line serial], 3(3). Available:http://www.nova.edu/ssss/QR/QR3-3/tellis2.html

[11] Current procedural terminology (Standard ed. : 1998) *Current procedural terminology : CPT / American Medical Association*. Chicago, IL : American Medical Association, c1998-

[12] Powers, Daniel A., and Yu Xie. *Statistical Methods for Categorical Data Analysis*. Orlando: Academic Press, 2000.

[13] "Probability and Statistics for Engineers and Scientists", 2nd ed., Walpole and Myers, MacMillian, 1978.

[14] Montgomery, D. (2001). *Design and analysis of experiments*. 5th ed. New York: John Wiley & Sons, Inc..

[15] Earnshaw, Stephanie R., and Susan L. Dennett. "Integer/Linear Mathematical Programming Models." *Pharmocoeconomics* 21.12 (2003): 839-851.

[16] Grunow, Martin, Hans-Otto Gunther, and Gang Yang. "Development of a Decision Support Model for Scheduling Clinical Studies and Assigning Medical Personnel." *Health Care Management Science* 7 (2004): 305-317.

[17] Jaumard, Brigitte, Frederic Semet, and Tsevi Vovor. "A generalized linear programming model for nurse scheduling." *European Journal of Operational Research* 107 (1998): 1-18.

[18] Crain, Marion. "The Transformation of the Professional Workforce." *Chicago-Kent Law Review* 79 (2004): 543-615.

[19] Knapp, Richard M. "Complexity and Uncertainty in Financing Graduate Medical Education." *Academic Medicine* 77 (2002): 1076-1083.

[20] Abrass, Christine K., Ruth Ballweg, Margaret Gilshannon, and John B. Coombs. "A Process for Reducing Workload and Enhancing Residents' Education at an Academic Medical Center." *Academic Medicine* 76 (2001): 798-805.

[21] Buchanan, James, and Mario R. Dal Poz. "Skill mix in the health care workforce: Reviewing the evidence." *Bulletin of the World Health Organization* 80.7 (2002): 575-580.

[22] "Industry-Occupation Employment Matrix: Occupation Report." *U.S. Department of Labor Bureau of Labor Statistics*. U.S. Bureau of Labor Statistics. 16 May. 2005 <http://data.bls.gov/oep/servlet/oep.nioem.servlet.ActionServlet?Action=empior&MultipleSelect=000000&Sort=ws_emp_b&StartItem=0&Resort=

No&ResortButton=No&Base=2002&Proj=2012&SingleSelect=29107104 11&Type=Occupation&Number=10>.

[23] "Engineering the Delivery of Health Care: Priorities for Application and Research ." *Engineering the Delivery of Health Care: Priorities for Application and Research*_. National Academy of Engineering of the National Academies. 16 May. 2005 <http://www.nae.edu/nae/naepcms.nsf/weblinks/MKEZ-5LWQVN?OpenDocument>.

[24] Fink, Arlene. *The Survey Handbook*. 2nd ed. Thousand Oaks: Sage Publications, 2003

[25] Fink, Arlene. *How to Design Survey Studies*. 2nd ed. Thousand Oaks: Sage Publications, 2003.

[26] Grant, Stan, Tom Aitchison, Esther Henderson, and Jim Christie. "A Comparison of the Reproducibility and the Sensitivity to Change of Visual Analogue Scales, Borg Scales, and Likert Scales in Normal Subjects During Submaximal Exercise*." *Chest* 116.5 (1999): 1208-1217.

[27] Guyatt, G H, M Townsend, L B Berman, and J L Keller. "A comparison of Likert and visual analogue scales for measuring change in function.." *Journal of Chronic Diseases* 40 (1987): 1129-1133.

[28] Torrance, George W., David Feeny, and William Furlong. "Visual Analog Scales: Do They Have a Role in the Measurement of Preferences for Health States?." *Medical Decision Making* 21.4 (2001): 329-334.

APPENDICES

Appendix A Residency Program Director Survey Instrument

Introduction / Instructions

Dear Program Directors:

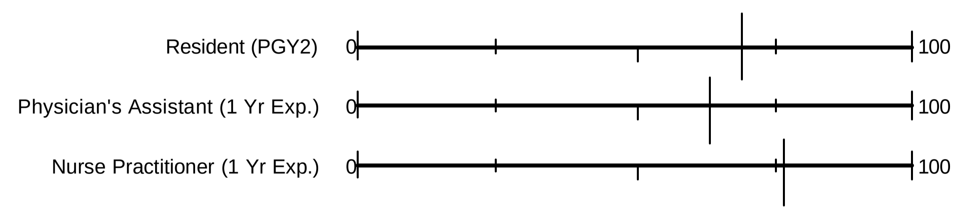
The following 16 questions seek your assessment of the performance of three classes of healthcare providers relative to the performance of an attending physician. The healthcare providers are: 1) Residents (PGY2), 2) Physician's Assistants (1 Yr Experience), and 3) Nurse Practitioners (1 Yr Experience).

Fifteen (15) of the questions will ask for your assessment based on three variables (Time, Quality, Independence) for five specific tasks directly related to E/M codes.

Please indicate your responses by placing a vertical mark along the visual analog scale adjacent to the healthcare provider in question.

For example: Question 1

The amount of TIME required to conduct an INPATIENT ADMISSION?
[E/M Codes: 99221 - 99223, 99234 - 99236]



The above would indicate responses of 68 for a Resident, 62 for a Physician's Assistant, and 72 for a Nurse Practitioner on a 0 to 100 scale with an Attending Physician rated at 100.

The final question is open-ended, and thus requires a written response.

THANK YOU FOR YOUR VITAL CONTRIBUTION TO THIS STUDY.

Appendix A (Continued)

How would you rate the proficiency of the following healthcare providers (relative to an attending physician as 100% proficient) in:

Q1. The amount of TIME required to conduct an INPATIENT ADMISSION?
[E/M Codes: 99221 - 99223, 99234 - 99236]

Resident (PGY2) 0 — 100

Physician's Assistant (1 Yr Exp.) 0 — 100

Nurse Practitioner (1 Yr Exp.) 0 — 100

Q2. The amount of TIME required to conduct an INPATIENT CONSULTATION?
[E/M Codes: 99251 - 99255]

Resident (PGY2) 0 — 100

Physician's Assistant (1 Yr Exp.) 0 — 100

Nurse Practitioner (1 Yr Exp.) 0 — 100

Q3. The amount of TIME required to perform EMERGENCY DEPARTMENT SERVICES?
[E/M Codes: 99281 - 99285]

Resident (PGY2) 0 — 100

Physician's Assistant (1 Yr Exp.) 0 — 100

Nurse Practitioner (1 Yr Exp.) 0 — 100

Q4. The amount of TIME required to perform CRITICAL CARE EVALUATION?
[E/M Code: 99291]

Resident (PGY2) 0 — 100

Physician's Assistant (1 Yr Exp.) 0 — 100

Nurse Practitioner (1 Yr Exp.) 0 — 100

Appendix A (Continued)

How would you rate the proficiency of the following healthcare providers (relative to an attending physician as 100% proficient) in:

Q5. The amount of TIME required to perform DISCHARGE DAY MANAGEMENT [HOSPITAL DISCHARGE SERVICES]?
[E/M Code: 99238]

Resident (PGY2) 0 ——— 100

Physician's Assistant (1 Yr Exp.) 0 ——— 100

Nurse Practitioner (1 Yr Exp.) 0 ——— 100

Q6. The QUALITY of the outcome in conducting an INPATIENT ADMISSION?
[E/M Codes: 99221 - 99223, 99234 - 99236]

Resident (PGY2) 0 ——— 100

Physician's Assistant (1 Yr Exp.) 0 ——— 100

Nurse Practitioner (1 Yr Exp.) 0 ——— 100

Q7. The QUALITY of the outcome in conducting an INPATIENT CONSULTATION?
[E/M Codes: 99251 - 99255]

Resident (PGY2) 0 ——— 100

Physician's Assistant (1 Yr Exp.) 0 ——— 100

Nurse Practitioner (1 Yr Exp.) 0 ——— 100

Appendix A (Continued)

How would you rate the proficiency of the following healthcare providers (relative to an attending physician as 100% proficient) in:

Q8. The QUALITY of the outcome in performing EMERGENCY DEPARTMENT SERVICES?

[E/M Codes: 99281 - 99285]

Resident (PGY2) 0 ——— 100

Physician's Assistant (1 Yr Exp.) 0 ——— 100

Nurse Practitioner (1 Yr Exp.) 0 ——— 100

Q9. The QUALITY of the outcome in performing CRITICAL CARE EVALUATION?

[E/M Code: 99291]

Resident (PGY2) 0 ——— 100

Physician's Assistant (1 Yr Exp.) 0 ——— 100

Nurse Practitioner (1 Yr Exp.) 0 ——— 100

Q10. The QUALITY of the outcome in performing DISCHARGE DAY MANAGEMENT [HOSPITAL DISCHARGE SERVICES]?

[E/M Code: 99238]

Resident (PGY2) 0 ——— 100

Physician's Assistant (1 Yr Exp.) 0 ——— 100

Nurse Practitioner (1 Yr Exp.) 0 ——— 100

Appendix A (Continued)

How would you rate the proficiency of the following healthcare providers (relative to an attending physician as 100% proficient) in:

Q11. The level of INDEPENDENCE (capability of working without supervision) demonstrated in conducting an INPATIENT ADMISSION?
[E/M Codes: 99221 - 99223, 99234 - 99236]

Resident (PGY2) 0 ———— 100

Physician's Assistant (1 Yr Exp.) 0 ———— 100

Nurse Practitioner (1 Yr Exp.) 0 ———— 100

Q12. The level of INDEPENDENCE (capability of working without supervision) demonstrated in conducting an INPATIENT CONSULTATION?
[E/M Codes: 99251 - 99255]

Resident (PGY2) 0 ———— 100

Physician's Assistant (1 Yr Exp.) 0 ———— 100

Nurse Practitioner (1 Yr Exp.) 0 ———— 100

Q13. The level of INDEPENDENCE (capability of working without supervision) demonstrated in performing EMERGENCY DEPARTMENT SERVICES?
[E/M Codes: 99281 - 99285]

Resident (PGY2) 0 ———— 100

Physician's Assistant (1 Yr Exp.) 0 ———— 100

Nurse Practitioner (1 Yr Exp.) 0 ———— 100

Appendix A (Continued)

How would you rate the proficiency of the following healthcare providers (relative to an attending physician as 100% proficient) in:

Q14. The level of INDEPENDENCE (capability of working without supervision) demonstrated in performing CRITICAL CARE EVALUATION?
[E/M Code: 99291]

Resident (PGY2) 0 ———— 100

Physician's Assistant (1 Yr Exp.) 0 ———— 100

Nurse Practitioner (1 Yr Exp.) 0 ———— 100

Q15. The level of INDEPENDENCE (capability of working without supervision) demonstrated in performing DISCHARGE DAY MANAGEMENT [HOSPITAL DISCHARGE SERVICES]?
[E/M Code: 99238]

Resident (PGY2) 0 ———— 100

Physician's Assistant (1 Yr Exp.) 0 ———— 100

Nurse Practitioner (1 Yr Exp.) 0 ———— 100

Q16. What other factors do you believe can be used to distinguish between a resident, physician's assistant and nurse practioner (other than time, quality, and independence)?

Appendix B Linear Programming Model and Output

```
Healthcare System 1

max 73 R + 56 PA + 55 NP

subject to

50165 R + 83820 PA + 82550 NP <= 14200000

R <= 10

R <= 20

2840 R + 2320 PA + 2240 NP <= 24461

R >= 0
PA >= 0
NP >= 0

end

gin R
gin PA
gin NP
```

```
OBJECTIVE FUNCTION VALUE
      1)      622.0000

 VARIABLE        VALUE          REDUCED COST
       R        7.000000       -73.000000
      PA        1.000000       -56.000000
      NP        1.000000       -55.000000


     ROW   SLACK OR SURPLUS     DUAL PRICES
      2)    482475.000000         0.000000
      3)        53.000000         0.000000
      4)       409.000000         0.000000
      5)        21.000000         0.000000
      6)         7.000000         0.000000
      7)         1.000000         0.000000
      8)         1.000000         0.000000

 NO. ITERATIONS=      26
```

Appendix B (Continued)

```
Healthcare System 2

max 73 R + 56 PA + 55 NP

subject to

50165 R + 83820 PA + 82550 NP <= 5600000

R <= 4

R <= 8

2840 R + 2320 PA + 2240 NP <= 5322

R >= 0
PA >= 0
NP >= 0

end

gin R
gin PA
gin NP
```

```
OBJECTIVE FUNCTION VALUE
      1)      129.0000

 VARIABLE        VALUE          REDUCED COST
       R         1.000000        -73.000000
      PA         1.000000        -56.000000
      NP         0.000000        -55.000000


     ROW   SLACK OR SURPLUS     DUAL PRICES
      2)   5466015.000000         0.000000
      3)         3.000000         0.000000
      4)         7.000000         0.000000
      5)       162.000000         0.000000
      6)         1.000000         0.000000
      7)         1.000000         0.000000
      8)         0.000000         0.000000

 NO. ITERATIONS=      13
```

Appendix C Data Analysis of Residency Programs at USF College of Medicine

Univariate Analyses

Kruskal-Wallis Test: Proficiency versus Provider

```
Kruskal-Wallis Test on Proficiency

Provider    N    Median    Ave Rank        Z
1          75     67.00       136.4     3.81
2          75     63.00       102.2    -1.77
3          75     63.00       100.5    -2.04
Overall   225                 113.0

H = 14.55  DF = 2  P = 0.001
H = 14.59  DF = 2  P = 0.001 (adjusted for ties)
```

Kruskal-Wallis Test: Proficiency versus Task

```
Kruskal-Wallis Test on Proficiency

Task        N    Median    Ave Rank        Z
1          45     63.00       108.7    -0.50
2          45     63.00       113.5     0.06
3          45     67.00       129.0     1.84
4          45     54.00        86.8    -3.01
5          45     67.00       127.0     1.61
Overall   225                 113.0

H = 12.28  DF = 4  P = 0.015
H = 12.31  DF = 4  P = 0.015 (adjusted for ties)
```

Kruskal-Wallis Test: Proficiency versus Dimension

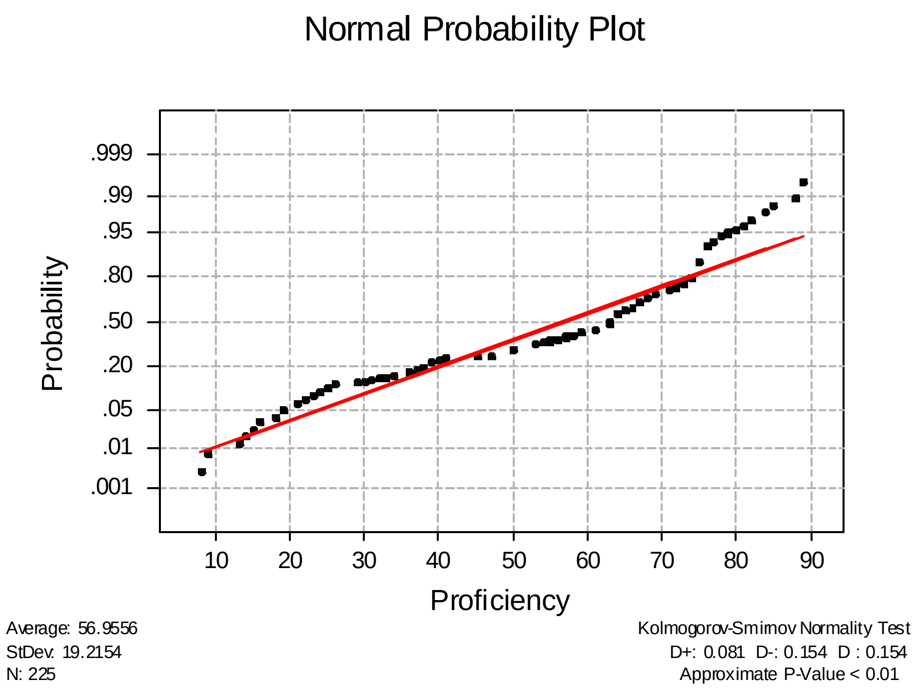
```
Kruskal-Wallis Test on Proficiency

Dimension   N    Median    Ave Rank        Z
1          75     58.00       101.0    -1.96
2          75     64.00       126.2     2.14
3          75     63.00       111.9    -0.18
Overall   225                 113.0

H = 5.65  DF = 2  P = 0.059
H = 5.66  DF = 2  P = 0.059 (adjusted for ties)
```

Appendix C (Continued)

Multivariate Analyses



General Linear Model: Proficiency versus Provider, Task, Dimension (Residency Programs-COMED)

Factor	Type	Levels	Values
Provider	fixed	3	1 2 3
Task	fixed	5	1 2 3 4 5
Dimension	fixed	3	1 2 3

Analysis of Variance for Proficiency, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Provider	2	4641.3	4641.3	2320.7	6.37	0.002
Task	4	3756.1	3756.1	939.0	2.58	0.039
Dimension	2	2378.8	2378.8	1189.4	3.27	0.040
Provider*Task	8	3447.2	3447.2	430.9	1.18	0.312
Provider*Dimension	4	276.2	276.2	69.1	0.19	0.944
Task*Dimension	8	1748.3	1748.3	218.5	0.60	0.777
Provider*Task*Dimension	16	898.8	898.8	56.2	0.15	1.000
Error	180	65560.8	65560.8	364.2		
Total	224	82707.6				

Appendix C (Continued)

Tukey Pairwise Comparison for Provider

```
Tukey 95.0% Simultaneous Confidence Intervals
Response Variable Proficie
All Pairwise Comparisons among Levels of Provider

Provider = 1 subtracted from:

Provider     Lower    Center     Upper  -----+---------+---------+---------+-
2           -16.76    -9.400    -2.040   (----------*---------)
3           -17.21    -9.853    -2.493  (----------*---------)
                                        -----+---------+---------+---------+-
                                           -14.0      -7.0       0.0       7.0

Provider = 1 subtracted from:

Level        Difference       SE of             Adjusted
Provider       of Means  Difference   T-Value    P-Value
2                -9.400       3.117    -3.016     0.0082
3                -9.853       3.117    -3.162     0.0052

Provider = 2 subtracted from:

Level        Difference       SE of             Adjusted
Provider       of Means  Difference   T-Value    P-Value
3               -0.4533       3.117   -0.1455     0.9884
```

Appendix C (Continued)

Tukey Pairwise Comparison for Task

```
Tukey 95.0% Simultaneous Confidence Intervals
Response Variable Proficie
All Pairwise Comparisons among Levels of Task

Task = 1 subtracted from:

Task      Lower    Center     Upper  ---------+---------+---------+-------
2         -9.54     1.556    12.651            (--------*---------)
3         -6.38     4.711    15.807               (--------*--------)
4        -18.05    -6.956     4.140     (--------*--------)
5         -7.63     3.467    14.562              (--------*--------)
                                     ---------+---------+---------+-------
                                            -12         0        12

Task = 1 subtracted from:

Level    Difference       SE of              Adjusted
Task       of Means  Difference   T-Value     P-Value
2             1.556       4.023     0.387      0.9952
3             4.711       4.023     1.171      0.7679
4            -6.956       4.023    -1.729      0.4190
5             3.467       4.023     0.862      0.9105

Task = 2 subtracted from:

Level    Difference       SE of              Adjusted
Task       of Means  Difference   T-Value     P-Value
3             3.156       4.023     0.784      0.9349
4            -8.511       4.023    -2.115      0.2181
5             1.911       4.023     0.475      0.9895

Task = 3 subtracted from:

Level    Difference       SE of              Adjusted
Task       of Means  Difference   T-Value     P-Value
4            -11.67       4.023    -2.900      0.0338
5             -1.24       4.023    -0.309      0.9980

Task = 4 subtracted from:

Level    Difference       SE of              Adjusted
Task       of Means  Difference   T-Value     P-Value
5             10.42       4.023     2.590      0.0764
```

Appendix C (Continued)

Tukey Pairwise Comparison for Dimension

```
Tukey 95.0% Simultaneous Confidence Intervals
Response Variable Proficiency
All Pairwise Comparisons among Levels of Dimension

Dimension = 1 subtracted from:

Dimension     Lower    Center     Upper  -----+---------+---------+---------+-
2             0.600     7.960     15.32                 (--------*--------)
3            -3.614     3.747     11.11           (---------*--------)
                                         -----+---------+---------+---------+-
                                            -8.0       0.0       8.0      16.0


Dimension = 1 subtracted from:

Level         Difference       SE of             Adjusted
Dimension       of Means  Difference   T-Value    P-Value
2                  7.960       3.117     2.554     0.0307
3                  3.747       3.117     1.202     0.4534

Dimension = 2 subtracted from:

Level         Difference       SE of             Adjusted
Dimension       of Means  Difference   T-Value    P-Value
3                 -4.213       3.117    -1.352     0.3685
```

Appendix D Definitions and E/M Codes for Tasks/Procedures

Task	E/M Codes	Definition
Inpatient Admission	99221 – 99223 99234 – 99236	Initial hospital care, per day, for the evaluation and management of a patient requiring: • A detailed/comprehensive history • A detailed/comprehensive examination • Medical decision making of low/moderate/high complexity Observation or inpatient hospital care for the evaluation and management of a patient including admission and discharge on the same date requiring the above (bulleted list).
Inpatient Consultation	99251 – 99255	Initial inpatient consultation for a new or established patient requiring: • A problem-focused/detailed / comprehensive history • A problem-focused/detailed / comprehensive examination • Medical decision making of low/moderate/high complexity
Emergency Department Services	99281 – 99285	The provision of unscheduled episodic services to patients who present for immediate medical attention. Emergency department visit for the evaluation and maintenance of a patient requiring the above (bulleted list from Inpatient Consultation).

Appendix D (Continued)

Task	E/M Codes	Definition
Critical Care Evaluation	99291	Critical care is the direct delivery by a physician(s) of medical care for a critically ill or injured patient . . . the care of such patients involves decision making of high complexity to assess, manipulate, and support central nervous system failure, circulatory failure, shock-like conditions, renal, hepatic, metabolic, or respiratory failure, postoperative complications, overwhelming infection, or other vial system functions to treat single or multiple vital organ system failure or to prevent further deterioration.
Discharge Day Management	99238	The time spent for final discharge of a patient. This includes, as appropriate, final examination of the patient, discussion of the hospital stay, instructions for continuing care to all relevant caregivers, and preparation of discharge records, prescriptions, and referral forms.