OUTPATIENT IMAGING IN PRIMARY CARE

By

CHRISTOPHER L. SISTROM

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2010

To my wife, Brenda

ACKNOWLEDGMENTS

The research described in this dissertation was made possible through collaboration with numerous people at the Massachusetts General Hospital Physician Organization and Radiology Department. They include (but are not limited to) Jeffrey Weilburg, MD, Keith Dreyer, DO, PhD, Timothy Ferris, MD, George Leehan, and Markus Stout. I am very grateful for their help and support.

Also, this research would not have been possible without the meticulous and consistent work by Dr. Steve Atlas and his group to identify a stable cohort of patients and primary care doctors.

My colleagues and friends in the Department of Radiology at the University of Florida have been very kind in allowing me to spend enough time away from routine clinical, administrative, and educational duties to engage in a multi-year effort to learn several new disciplines; epidemiology, public health, health services research. Without considerable support and encouragement on their part, none of this would have been possible. They include (but are not limited to) Anthony Mancuso, MD, Patricia Abbitt, MD, Jonathon Williams, MD, Melinda Chitty, Anne D'Amico, and Meryll Frost.

Niccie McKay, PhD has been tireless and infinitely patient in several roles including mentor, teacher, editor, and chair of my dissertation committee. Cyndi Garvan, PhD gave me a crash course in hierarchical modeling and was the external member of my committee. Jeffrey Harman, PhD, has been instrumental in bringing a methodological rigor to my efforts and Chris Harle, PhD honored me by inaugurating his faculty career in serving on my dissertation committee.

My wife Brenda has been a wonderful partner during what may have seemed at times like a decade long mid-life crisis. I very much appreciate her support and

forbearance and am glad to know that she will be around as I decide ‘what I want to be when I grow up’.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

OUTPATIENT IMAGING IN PRIMARY CARE

By

Christopher L. Sistrom

August 2010

Chair: Niccie McKay
Major: Health Services Research

Diagnostic imaging comprises a rapidly growing portion of health care dollars spent in the U.S. Additionally, imaging tests that use X-Rays (including C.T. scans) now contribute half of the entire radiation dose to the population; having increased from 15% in 1980. Like other medical services, diagnostic imaging tests are utilized in some states and cities much more frequently than in others, even after controlling factors like age, gender, and illness burden. This marked variability in how frequently and for what reasons that imaging tests are done extends down to the level of individual doctors. Understanding the causes and effects of these sorts of differences in how health care is delivered (including imaging tests) is one of the core parts of health services research.

The fact that patients in a stable relationship with a primary care doctor have better health outcomes and consume less health services overall has prompted efforts to increase the supply of primary care doctors and encourage patients to seek care in a so-called ‘medical home’ setting. One of the key roles of the primary care doctor is as a ‘gatekeeper’ for expensive tests and procedures as well as referral to specialists. This study looks at a large group of patients (about 85K) being taken care of by 148 primary care doctors to whom they have demonstrated ‘loyalty’ over three years based on the

pattern of office visits. This so called ‘loyalty cohort’ is useful as a representative sample of other patients and doctors functioning in the ideal doctor-patient relationship.

The study seeks to answer questions about how various factors in the patients (including clinical activity) relate to the amount of imaging tests that their primary care doctor orders for them over two years. This can be useful in comparing the use of imaging tests (e.g., images for 100 patients in a year) between primary care doctors by helping to adjust for differences in each doctor’s mixture of patients. Such risk-adjusted utilization profiles help to understand variability in practice style and resource use between doctors for purely scholarly interest as well as more practical uses by entities actually paying for the services (e.g., insurance companies, employers, government health programs).

A total of 35 pieces of information about each patient (not including their name or other identifiers) and 7 characteristics of each doctor were gathered from 6 different sources of data used routinely during patient care in the practices being studied. This information included complete listing of about ¼ million diagnostic imaging tests. Of these, about 60K were ordered by the patient’s own (loyal) primary care doctor and were counted by patient to form the variable of interest (outcome). Statistical relationships between all the other patient and doctor factors with the number of imaging tests were analyzed.

The results demonstrated that older female patients had more imaging as did those who had many medical problems listed in the clinical record system. Also, patients who visited doctors, were admitted to the hospital, or seen in the emergency room frequently had more imaging. Doctor factors associated with a greater tendency to

order imaging tests were less experience, female gender, and having a medium size practice (500-1000 patients). A special statistical technique that accounts for all patient factors allowed creation of ‘profiles’ scoring each of the 148 doctors on their general tendency to order imaging tests on the ‘average’ patient and how many more tests were ordered on patients with greater comparative ‘need’ for diagnostic imaging.

CHAPTER 1
INTRODUCTION

Background and Significance

Medical technology is often cited as a major driver of health care costs in the U.S., with diagnostic imaging being a 'poster child' of this trend. Advanced imaging, ordered during ambulatory care, is not only costly, but often leads to a 'cascade' of further tests and interventions (Deyo 2002, Mold and Stein 1986, Verrilli and Welch 1996). There is ample evidence of marked variation in utilization of imaging services in outpatient care at multiple levels of aggregation, from international down to individual groups within a large practice (Burkhardt and Sunshine 1996, Couchman *et al.* 2005, Goel *et al.* 1997, Hartley *et al.* 1987, Katz *et al.* 1996, Lysdahl and Borretzen 2007). This variation in utilization suggests that a substantial portion of diagnostic imaging may be unnecessary. Patients receiving such studies may be needlessly exposed to radiation. Findings of uncertain clinical meaning may prompt more imaging and costs may be increased.

Examining the ways in which primary care physicians utilize imaging, and using those insights to inform and enhance the consistency of their choices, can improve health care quality and increase cost effectiveness. Thus, the event of a primary care doctor ordering an imaging test is common and expensive enough to warrant study by itself. Also, it represents a very fruitful paradigm for understanding medical decision making and much of the variation in downstream discretionary health care utilization (Parchman 1995). The interesting events that occur in primary care settings are 'upstream' from major procedures and consist of referral to specialists and diagnostic testing. These discretionary decisions made during routine office visits to general

practitioners have substantial impact on population health and the overall cost of health care (Sirovich *et al.* 2008).

Specific Aims

- Quantify outpatient diagnostic imaging utilization over two years by primary care doctors caring for a cohort of patients that regularly attend clinic (i.e., are loyal)
- Collect and characterize demographic, contextual, and clinical factors for all patients in the cohort
- Collect and characterize demographic and practice factors for the primary care physicians regularly caring for the same patients in outpatient clinics
- Determine the relationships between patient, doctor, and clinic factors and the probability that patients had at least one imaging test during the two year study period (any use)
- Determine the relationships between patient, doctor, and clinic factors and the number of examinations performed in patients with at least one imaging test (intensity of use)
- Develop a model based risk adjustment method for imaging utilization (any use and intensity of use) producing an 'expected' amount of imaging given known patient and doctor factors (imaging propensity)
- Estimate and partition the variability in imaging utilization using a hierarchical method which takes into account patient's risk adjusted imaging propensity which is in turn influenced by each doctor's tendency to use imaging in their practice

Summary of Study

The basic unit of interest for this study is a patient who regularly visits a primary care physician and the time frame is two years. The main outcome is a measure of the amount of outpatient imaging performed in the period of study that was ordered by the same primary care physician on the patient in question. Specifying and quantifying the 'amount' of imaging is a non-trivial task because we wish to count all non-invasive diagnostic studies of any modality (e.g., CT, MRI, XRAY, Ultrasound, Nuclear Medicine,

PET). Fortunately, the Radiology Department and the host institution for the study share a longstanding commitment to a robust medical informatics infrastructure. Thus, complete and detailed records of all imaging events for the past 15 years are readily accessible. Also, to enhance medical management and clinical operations support, a full spectrum of clinical data are available for health services research such as this study. These informatics resources allow collecting a large amount of data on each patient that will be used to form explanatory and control variables for analyzing drivers, enablers, and inhibitors of outpatient primary care imaging utilization. Finally, using credentialing databases from the host institution and publicly available data from state licensing sources, provider level factors will be obtained for study.

After characterizing and quantifying the imaging utilization over two years of study, the tests performed on each patient will be attributed to them and cross-tabulated according to the patient's status (Inpatient, Emergency Room, Outpatient) when performed and the type of provider ordering the study (primary care vs specialist). Patient and provider level variables will be characterized, validated, and then individually evaluated in bivariate fashion with the main outcome (count of outpatient imaging tests ordered by patient's primary care doctor). Initial multivariable modeling will be performed using logistic regression on the whole data set (outcome=any imaging yes/no) followed by linear regression (with Poisson errors) on the non-zero observations. The joint contribution of the various patient, provider, and clinic factors to any use (logistic regression) and intensity of use (Poisson regression) will be inferred from the odds ratios and coefficients respectively.

To prepare for hierarchical analysis, a zero-inflated Poisson regression will be estimated using all observations and only patient level factors to form a single continuous variable from the model predictions. This serves to summarize the multiple patient level factors into a single number representing risk adjusted expected imaging or 'propensity' for each patient. A two-level hierarchical model with patients at the first level nested within the 148 providers at the second level will be specified and fitted to the observed imaging counts for all patients. Only two independent variables will be included: the propensity variable and a unique ID number for each provider. By specifying a model that defines provider level intercepts (mean imaging use) and slopes (response to imaging propensity) two unique characteristics can be estimated for each provider and compared with each other. In addition, variance components computed for each level in the hierarchical structure will serve to partition variation in imaging utilization between patient's imaging propensity, provider's general tendency to image, and provider's response to clinical factors in their patients in the amount of imaging they order.

Contribution to Literature

This research will contribute to the literature about primary care imaging utilization in several ways. First of all, the study population of providers and patients is large, comprehensive, and unique (i.e., all primary care doctors and their loyal patients at a large metropolitan academic health center). A previously validated loyalty cohort methodology (detailed in setting, data, and variables section) identifies patients, doctors, and clinics in ongoing and stable relationships with each other (i.e., usual source of care or 'medical home'). The available data about imaging utilization and patient health status is complete and highly detailed, having come directly from primary

sources (i.e., electronic medical records and clinical radiology information systems). For example, the requesting provider is recorded for each imaging test, which allows stratification of utilization by who ordered the examination (patient's primary care physician versus a specialist) and the patient care setting (outpatient, inpatient, emergency department).

Complete data are also available for all outpatient visits, hospital stays, and emergency room encounters. Therefore, the analytic data will contain a robust set of clinical activity and risk adjustment variables, which will be used to estimate the risk adjusted expectation (propensity) for imaging utilization to a high degree of accuracy and precision. Once patient level clinical factors have been accounted for, residual variation between providers will be quantified and partitioned in a way that should shed considerable light on the contribution of various contextual factors. When analysis of primary care outpatient imaging utilization is extended to provider profiling, proper specification of patient level risk-adjustment and hierarchical analysis at the provider level is crucial to fairly applying these medical management tools and this study will advance knowledge of these issues.

Policy Implications

In the ongoing debate about national health care policy the most divisive and vexing issues relate to unsustainable medical cost inflation. This is largely the result of increasing utilization of expensive, high-technology diagnostic and therapeutic interventions which occurs at the discretion of physicians to a substantial degree. The magnitude of this physician discretion over utilization as opposed to evidence-based clinical need will determine success of strategies to reduce costs targeted to physicians. These include education, point of care intervention, clinical decision support, and post

hoc profiling for ‘efficiency’. This study will provide estimates for the relative contributions of clinical need, physician style, and non-clinical patient factors with respect to utilization of outpatient diagnostic imaging in primary care. Assuming that clinical need remains as a significant and substantial driver of imaging utilization, this study provides insight into risk adjustment models which will be necessary for utilization management efforts going forward.

CHAPTER 2
BACKGROUND

This chapter will provide background material about medical imaging as a diagnostic test applied in routine outpatient primary care. Explanation of exactly what is meant by adult outpatient primary care will serve to further refine the setting for the research described in this dissertation. In addition to specifying exactly what procedures are and are not included in a definition of diagnostic imaging, the chapter will describe national trends toward rapidly increasing volume and cost of these services.

Definition of Diagnostic Imaging

A diagnostic imaging procedure (DIP) is defined as a discrete event with the following attributes. The ‘subject’ of this event is an individual in a provider-patient relationship with a medical practitioner. The practitioner initiates the event by means of an order for the DIP, and this can be verbal, written, or electronic. The individual submitting the order will be called the ordering provider (OP). Two other provider roles are required to complete the event and these are the performing provider (PP) and the interpreting provider (IP). The PP interacts directly with the patient using some kind of diagnostic imaging equipment to produce images which are subsequently interpreted by the IP who communicates the findings to the OP. Note that the while the three providers (OP, PP, and IP) are often separate individuals, a single person may perform all three roles. For example, an obstetrician may decide that a fetal ultrasound should be done on her patient, may personally perform the scan, and interpret the images in real-time from the video display all in a single step. Formal communication of the interpretation back to the OP typically includes some permanent documentation of the findings and interpretation, which is placed into the patient’s medical record.

In addition to the people involved, other attributes serve to describe a DIP. Although defined as a discrete event, a DIP occurs in a series of steps: ordering, performing, interpreting, and review of results. This has been referred to as the 'radiology round trip' to emphasize the complexity of the process (Thrall 2005b, Thrall 2005a, Thrall 2005c). In some settings, the time between these steps can be lengthy and quite variable. For most purposes, we define the procedure as having 'occurred' at a single point in time, when the PP finishes performing or 'completes' the examination. This is often referred to as the 'date of service' in medical record systems, billing applications and in claims data.

Two other attributes of a DIP must be articulated and these define the sort of equipment used to obtain the images (modality) and what part of the patient (body area) was examined. The term 'modality' refers to the physical nature of the process used to create images of the patient's body and is useful as a primary means of categorizing the equipment. For example, 'radiography' uses invisible photons of high energy (X-Rays) directed at and through the patient from a fixed generating tube and detects the photons on a flat surface (film or digital plate). Computed Tomography (CT) also uses X-Rays but the generating tube spins around the patient along with a detector. The varying intensity of photons falling on the detector is combined with angular position to compute a tomographic (planar slice) image of the patient's inner structure. Magnetic resonance imaging (MRI) machines irradiate patients with sequences of radio waves from coils housed in a strong magnetic field. Very sensitive antennae detect weak 'echoes' of the radio waves emitted by hydrogen atoms within the patient's tissues and compute an image based on strength and frequency. Nuclear medicine (including positron emission

tomography=PET) involves injecting patients with a variety of radioactive ‘tracers’ which emit gamma ray photons or energetic particles as they undergo decay. These are detected outside the patient and processed to form a map or image of activity. Finally, ultrasound (US) produces pulses of high frequency sound from a specialized transducer in direct contact with the patient’s skin. The same transducer detects ‘echoes’ of those sound waves that differ in amplitude and timing depending on tissue characteristics and spatial location. These signals are processed into an image of anatomy immediately beneath the transducer which can be moved around to examine an entire region.

The final attribute assigned to a particular DIP details the anatomic regions of the patient (body area) that are exposed to the radiation, particles, or sound to produce images for subsequent interpretation. The nomenclature of body areas typically ‘imaged’ is fairly straightforward and includes: head, chest, abdomen, pelvis, spine, arms, and legs. More specialized examinations may cover particular anatomy such as coronary arteries, lungs, aorta, gallbladder, and so forth. When a modality is combined with a body area the result is a specific named DIP and a patient is said to have undergone (completed it) at the date/time (of service) that they were dismissed from the testing facility and the images became available for interpretation.

An example will help to clarify these definitions. Mary Jones (an adult) makes an appointment with her internist (Dr. Smith) and during the visit, complains of frequent and increasing headaches. Dr. Smith may decide to order a DIP to exclude the possibility of a structural lesion (e.g., a mass in the brain) before treating her with drugs to relieve the pain. In this case, Dr. Smith (OP) will order a test that evaluates the brain (body area = head) and has several options about modality including X-Ray, CT, MRI, Nuclear

Medicine and even ultrasound. Dr. Smith's decision as to which of these modalities to order relies on her personal assessment of the relative appropriateness for Ms. Jones. Assuming that they opt for MRI, Dr. Smith orders an MRI of the head/brain for Ms. Jones to be performed sometime in the future. Dr. Smith's order is transmitted in some fashion (computer order, phone, fax, paper prescription carried by Ms. Jones) to an imaging facility (often hospital based) that offers MRI scanning for outpatients. When Ms. Jones keeps her appointment, she is brought into the MRI scanning suite and asked to lie down on a movable couch which carries her into the actual MRI machine. She is instructed to breathe quietly and hold still while the technologists (PP) execute a pre-programmed protocol that directs the machine to acquire images of Ms. Jones' brain over the next several minutes. Upon completion of these imaging 'sequences' and a brief observation period, Ms. Jones is released to return home. At the same time, the PP (technologist) executes commands to 'complete' the examination at which time the images (computer files) are transmitted to storage media so as to be ready for download and interpretation and/or review. This sequence of events forms a single unit of outpatient imaging utilization and may be described by saying: Dr. Smith (OP) requested an MRI of the head on Ms. Jones (the patient) during an outpatient visit and it was completed on (the date of service) by (PP) to be interpreted by (IP) with the report to be sent back to Dr. Smith (OP) for review.

The next section deals with trends in the U.S. of rapidly increasing volumes of imaging utilization. To set the stage for this, consider the typical charges and reimbursements that might be submitted and paid, respectively, for Ms. Jones' MRI of the head. It is important to note that there are two separate 'billable' imaging events

arising from this sequence. The first is the encounter at the imaging facility where the MRI images of Ms. Jones' head were made. The outpatient imaging facility will bill Ms. Jones and/or her insurance carrier for the 'technical charge' which will likely be well over $1000.00. The second is the bill rendered by the physician who actually interprets the images, and this 'professional charge' will approach $500.00. The actual dollar amounts of reimbursement received will depend on the payer, with different payers reimbursing different amounts for the same service.

Of course, Dr. Smith will also submit a bill for the office visit during which she ordered the scan on Ms. Jones. However, this is not directly attributed to the imaging test and is not counted as part of its cost. It should also be noted that in an increasing number of cases, the MRI machine might actually be owned (or leased) by Dr. Smith and sometimes she might interpret the images herself. Under this scenario, Dr. Smith may directly bill for, or receive through more indirect means, most or all of the revenue generated by the technical and interpretation charges. This practice (called self-referral of imaging) is controversial among physicians and is targeted as a driver of increasingly burdensome costs in the U.S. by government and private payers.

Imaging Utilization and Costs

No other branch of medical technology has experienced the explosive growth in volume and variety of available services that radiology has during the past two decades. The medical care industry in the U.S. has purchased and installed advanced imaging equipment at an astounding rate, outpacing all other countries. Figure 2-1 illustrates this trend in terms of number of imaging procedures (Medicare Payment Advisory Commission 2005). Given that the current population of the U.S. is just now reaching 300M, this translates into about 1.4 imaging tests per person year. The most dynamic

growth has occurred in CT scanning, with a steady increase in capability and indications for use occurring over the last 30 years.

As the number of imaging procedures (many of which are CT scans) has increased, the cumulative effective radiation dose to the average American has nearly doubled from 3.6 milli-Sieverts (mSv) in 1980 to 6.2 mSv in 2006 (National Council on Radiation Protection & Measurements 2009). Several high profile articles in the past couple of years have raised concerns about a small but significant population risk for subsequent cancers induced by ionizing radiation delivered during medical imaging procedures (Brenner and Hall 2007, Fazel *et al.* 2009, Berrington de Gonzalez *et al.* 2009, Nyweide *et al.* 2009, Smith-Bindman *et al.* 2009).

Over the period 1985-1990, established technologies, such as CT, continued to grow in volume for Medicare. At the same time, the new technology of MRI exploded in terms of utilization with a 372% increase in national procedural volume for Medicare (Boutwell and Mitchell 1993). Imaging costs to the Medicare system in the past two decades rose much more rapidly than any other component and now comprise at least 14% of total Part B expenditures for physician services as specified in a report by the Medical Payment Advisory Commission (MEDPAC) to the U.S. Congress (Medicare Payment Advisory Commission 2003). Imaging costs grew by approximately 10% per year during the period covered by the report (1999-2002) compared with average growth of 3.3% per year for all physician services.

Testifying before Congress in 2006, Glenn Hackbarth of MEDPAC amplified and extended prior reports and testimony (Medicare Payment Advisory Commission 2006). He presented 1999-2004 data showing growth in Medicare claims for diagnostic

imaging as being the highest of all services at 62%. Furthermore, growth was especially high in emerging modalities (up to 140%) with even established technologies, like head CT, outpacing general growth at 43%. Figures 2-2 and 2-3 (exhibits in Mr. Hackbarth's testimony) illustrate these points and have been widely reproduced.

Primary Care Setting

This research examines imaging in outpatient adult primary care, which is defined as being rendered by doctors trained in internal medicine, family practice, and general practice. This is the classic paradigm of clinical decision-making, in which patients present with signs, symptoms, known diagnoses, or physical abnormalities that generate moderate probability of one or more treatable conditions. Much of the literature about the utilization of health services by primary care providers deals with one of two possible responses to this situation: diagnostic testing or referral to specialists.

This study will not consider imaging utilization that occurs as part of disease screening programs, including mammography, CT colonography, cardiac calcium detection with CT, lung cancer screening with CT, among several others. While important for population health and public policy, fundamental differences exist between imaging for screening and imaging for diagnostic or prognostic purposes. For example, once a universal population screening strategy has been adopted, policy-makers are primarily concerned with under-utilization in the target group. Conversely, in the case of diagnostic imaging, payers and regulators (at least in the U.S.) concentrate almost exclusively on problems of over-utilization.

Additionally, this work will not attempt to analyze utilization of imaging that occurs during inpatient care or imaging tests that are ordered during work-up of patients in emergency and urgent care settings due to the differing nature of medical decision-

making during inpatient and urgent care encounters compared with primary care. Further, the study does not examine imaging utilization directed by surgeons, oncologists or other specialists, because patients seen by doctors in these fields have unique, disparate, and complex medical problems and co-morbidities. Finally, imaging for children directed by pediatric providers will not be considered as it is quite rare in primary care settings and quickly devolves into specialty-oriented utilization after identification of congenital abnormalities, childhood cancer, or other serious problems.

Imaging as Diagnostic Testing

Imaging examinations that are interpreted by someone other than the ordering doctor represent a hybrid of diagnostic test and specialist referral. This is because imaging tests (especially complex ones like CT and MR) are perceived as consultations by patients as well as by ordering physicians. In contrast, clinical laboratory tests (blood chemistry, hematology, microbiology, and so forth) are ‘reported’ rather than ‘interpreted’ and this distinction is crucial. Non-imaging diagnostic test results are generally reported in the form of numbers or simple fact assertions and are often produced by automated methods with minimal analytic input by the rendering personnel. Examples of such tests include blood chemistry, antigen/antibody assay, serum drug levels, and urinalysis. With these, the ordering physician must synthesize an ‘interpretation’ about whether or not the result is normal and then decide relevance to the clinical question. On the other hand, radiologists produce interpretative documents that reach provisional conclusions about the probability of relevant clinical conditions (or at least classes of disease). Sometimes these reports contain recommendations for further follow up, clinical correlation, or even treatment and thus also function as consultation notes.

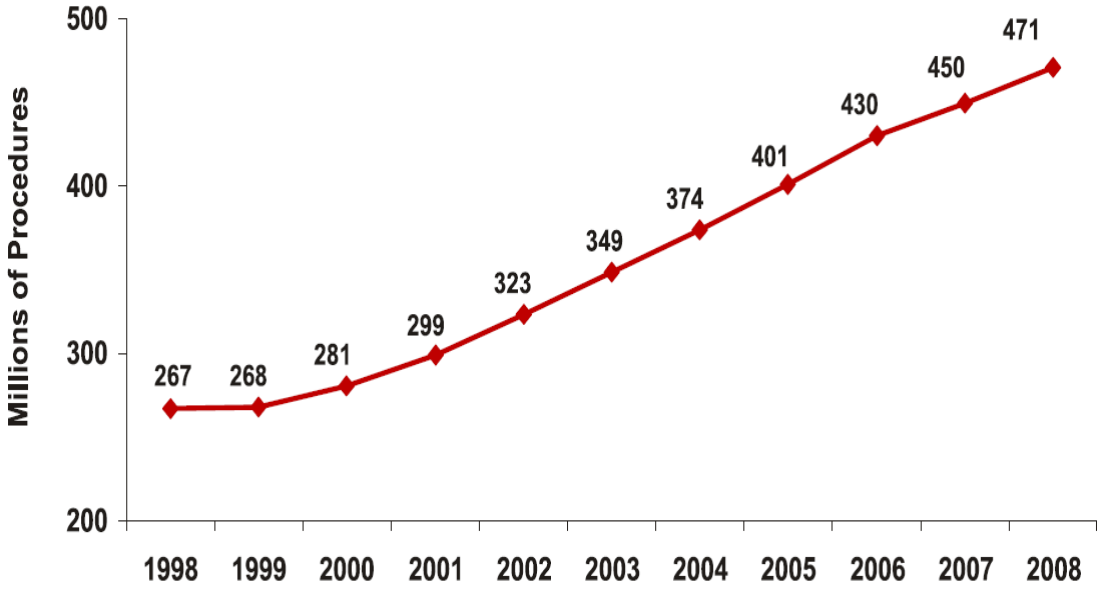

Figure 2-1. Total number of imaging studies by year in the U.S.

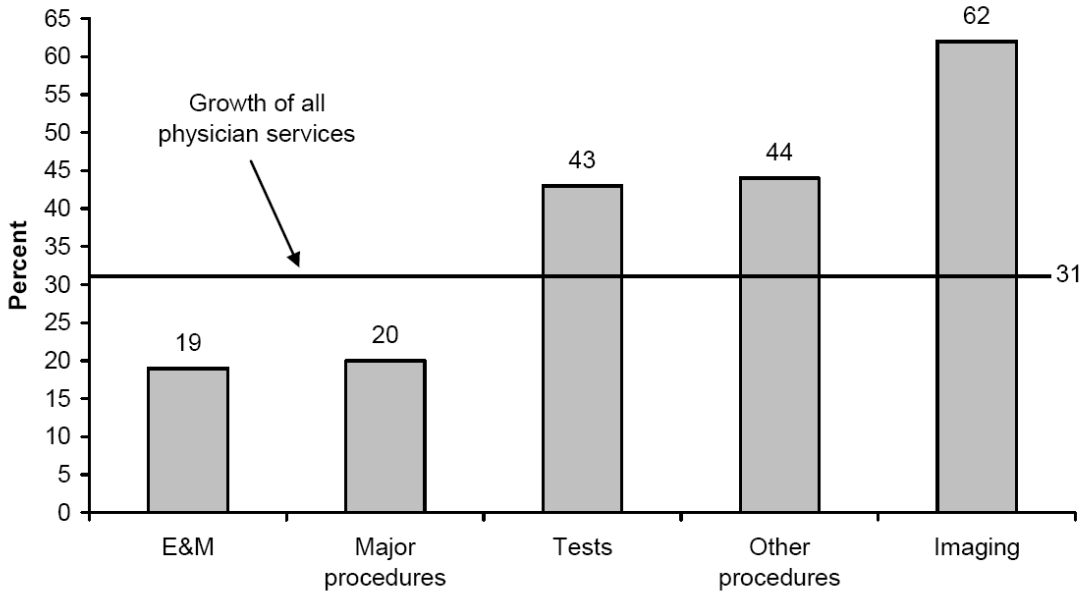

Figure 2-2. Imaging shows highest cumulative growth in services per beneficiary (1999-2004).

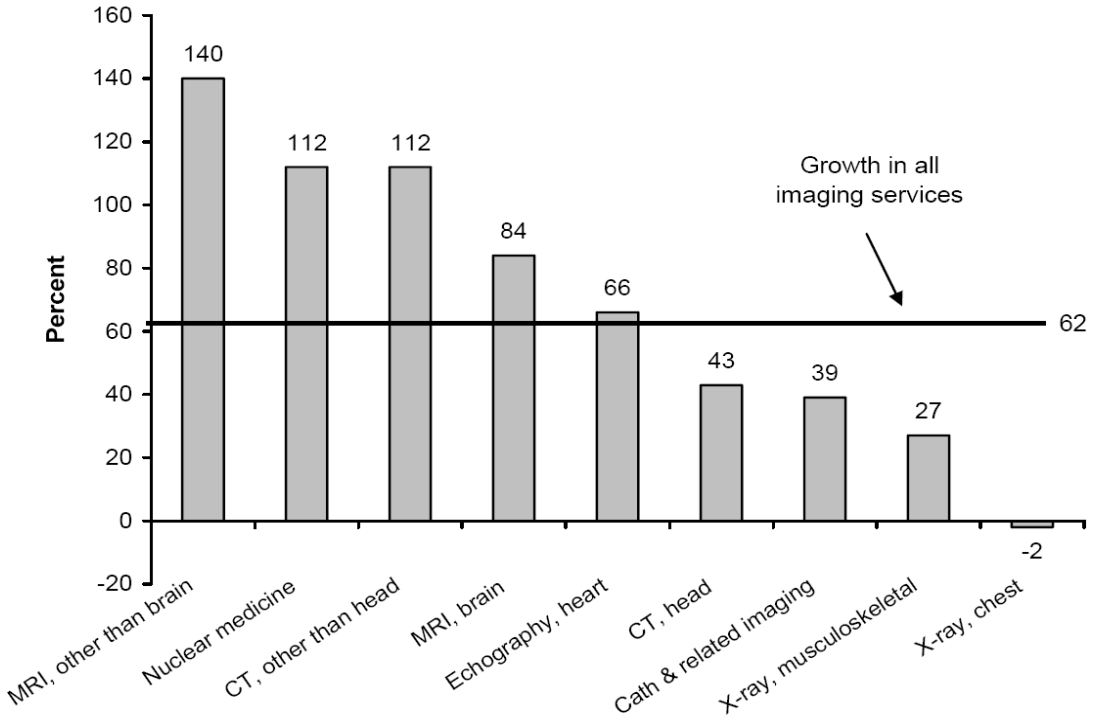

Figure 2-3. Cumulative growth in imaging volume varies by type (1999-2004).

CHAPTER 3
LITERATURE REVIEW

This chapter will summarize relevant publications that inform a conceptual model of imaging utilization in outpatient primary care. A general overview of theory and empiric research about small area variation in patterns of health service utilization will show that clinical uncertainty, physician practice style, patient preferences, and economic factors all play roles. An important contributor to variation in use of diagnostic imaging is clinical uncertainty. In fact, imaging tests are doubly subject to variation in utilization related to clinical uncertainty because, by definition, their main function is to reduce it. Thus, ambiguity about whether to observe, test, or treat in specific clinical situations is multiplied by uncertainty about which, if any, imaging test to order. In addition to how best to diagnose and treat their patients, doctors are also concerned about income, leisure time, satisfaction with practice, and mitigation of malpractice risk. These considerations influence utilization of imaging and their effect is magnified in the presence of clinical uncertainty. Finally, patients bring a complex mixture of factors to decisions about diagnostic imaging.

The Andersen Model

Any examination of health services utilization must consider the Andersen behavioral model as an organizing framework. The Andersen model seeks to explain health care utilization in an entire community, many members of whom do not seek medical care at all. Andersen reviewed his model 25 years after it was developed and described it mostly in terms of access to health care (Andersen 1995). The Andersen model is primarily applied to factors that determine the seeking of health care services in whole populations rather than what happens during encounters with providers (e.g.,

referral for diagnostic imaging). In a meta-analysis of papers using the Andersen model, only 2 of 139 examined provider characteristics such as specialty, experience, and gender (Phillips *et al.* 1998).

Thus the Andersen model is not directly applicable here because in looking at non-screening imaging utilization in primary care, the denominator is patients who already have an active relationship with a doctor. Like prescription drugs, one cannot undergo a diagnostic imaging test without a prescription or referral from a health care provider. In Folland's comprehensive review of variations in the use of medical care, this dichotomy in overall utilization is labeled as 'first occurrence' versus 'intensity' (Folland and Stano 1989). Much of the Andersen model deals with 'first occurrence' whereas imaging utilization falls under the 'intensity' concept. Therefore, the majority of this literature review will focus on works that inform the intensity of outpatient imaging utilization in an existing patient-provider relationship.

That being said, it will become evident from the empirical distribution of outpatient imaging examined in this study, that utilization also seems to have a two-part structure. These may be called *any use* and *amount* of--non-zero--use to account for a substantial fraction (over half) of patients with no imaging at all during the two years of data collection. However, in this study, even patients with no diagnostic imaging have visited their primary care doctor regularly and a differently specified conceptual model will be articulated below that is specific to outpatient imaging this special setting.

Small Area Variation

In an ideal world, the intensity and mixture of imaging would be appropriate to each patient's clinical situation regardless of any contextual differences among providers or patients. The counterfactual goes like this: Consider a situation where all

doctors and patients in a particular setting have access to identical and comprehensive evidence about imaging appropriateness (Sistrom 2009). Further, physicians act purely as agents and patients make rational choices based on maximizing longevity and well-being. Finally, economic considerations are uniform across all doctors and patients (e.g., a single payer universal insurance program). In this idealized setting, the amount and types of imaging tests that patients underwent would be based solely on their clinical presentation including pre-existing conditions and new symptoms, signs, and disease trajectory. This implies that, after fully accounting for differences in clinical presentation, the adjusted rate of imaging utilization would vary minimally over different levels of aggregation. That being said, there is little theoretical guidance as to exactly what this 'minimal' or 'natural' variation would entail in terms of intensity and mixture of imaging or any other health service (Cain and Diehr 1992, Diehr *et al.* 1990).

Empirical investigations have identified considerable variation in the use of medical services. The large body of literature documenting this phenomenon is referred to as research on "small area variations". The seminal paper on this topic, which was published by Wennberg and Gittelsohn in 1973, compared utilization and expenditures among 13 hospital service areas in Vermont during 1969 and found large variations among them (Wennberg and Gittelsohn 1973). For example, appendectomy rates varied from 10 per 10,000 persons to 32 per 10,000 persons. Since then, numerous articles, monographs and texts have examined geographic variations at the international, state, regional, and local levels for a wide variety of types of medical services (Health Services Research Group 1992, Folland and Stano 1990, Pekoz *et al.* 2003, Stano 1991, Stano 1993, Wennberg 1993). Wennberg's legacy is perhaps best

embodied in the Dartmouth Atlas of Healthcare which aggregates and analyzes Medicare claims going back to at least 1990 through the most recent past year to characterize variations in health service use across the entire U.S. in over 300 carefully defined contiguous regions (Wennberg 2004b) (Wennberg 2004a). Much of the available small area variations data deals with broad categories of services or spending. For example, the Dartmouth Atlas provides use rates and costs per beneficiary for the category of diagnostic services which includes laboratory tests and imaging procedures. When more granular area utilization analyses have been undertaken, they often have examined countable and more 'high impact' events like hospitalization, surgery, and invasive diagnostic procedures (e.g., endoscopy and cardiac catheterization) (Leape *et al.* 1990, Chassin *et al.* 1987).

The policy concern arising from evidence of small area variations is that they persist even after controlling for factors such as age, gender, and race (Health Services Research Group 1992, Blumberg 1987, Davis *et al.* 2000b, Wennberg 1987a, Wennberg 2002). Moreover, research has found little evidence that the geographic variations were correlated with differences in health status, disease-specific population health indicators, or other measures of health outcomes (Sirovich *et al.* 2006, Fisher *et al.* 2003b, Fisher *et al.* 2003a, Franks *et al.* 2000a, Brook and Lohr 1985). Thus, the lack of measurable population benefit in higher use areas combined with rising health care spending, has led policy makers to hope that health care cost reduction can be obtained without welfare loss by targeting upper outliers (Fisher *et al.* 2003a, Fisher *et al.* 2003b, Bodenheimer and Fernandez 2005, Stano 1993). With respect to imaging services in particular, a Medicare Payment Advisory Commission (MedPAC) report in

2006 noted that the geographic variation in utilization of imaging services (coefficient of variation of 28%) is exceeded only by non-imaging diagnostic tests (coefficient of variation of 30%) (Medicare Payment Advisory Commission 2006). An earlier study in 2001 found that there was a large variation by state in the number of diagnostic imaging studies per 1000 Medicare beneficiaries with 10th percentile being 3038 and the 90th percentile at 4573 (Bhargavan and Sunshine 2005).

As with other medical services that display small area variation, we must consider if inappropriate imaging accounts for the high-utilization areas and/or if poor access to services explains the low-use areas (Leape *et al.* 1990). In 2008, the General Accounting Office (GAO) published an extensive report about Medicare Part B imaging services that noted the rapid increase in spending and commented further on the extensive variation in utilization rates among states (Government Accountability Office (U.S.GAO) 2008). For example, by 2006, in-office imaging spending per beneficiary varied almost eight-fold across the states; from $62 in Vermont to $472 in Florida. The report goes on to state that "Given the magnitude of the differences in imaging use across geographic areas, variation is more likely due to differences in physician practice patterns rather than patient health status. Further concerns about the appropriateness of imaging use are raised by research on geographic variation showing that, in general, more health care services do not necessarily lead to improved outcomes" (page 21).

Just as scientists leverage observed variability in nature to pry into its inner workings, health policy makers target variability in utilization to better understand and control overall cost and quality. After characterizing and quantifying medical practice variation, focus naturally turns towards explaining it (Folland and Stano 1990, Folland

and Stano 1989). In ambulatory primary care, much of this attention is devoted to two processes; specialist referral (Franks *et al.* 2000a, Franks *et al.* 2000b, Franks *et al.* 1999) and diagnostic testing (Hartley *et al.* 1987, Epstein and McNeil 1987, Epstein and McNeil 1986, Epstein and McNeil 1985b, Epstein and McNeil 1985a). When making these decisions, providers and patients labor under compound uncertainty about the presence of disease(s) as well as the efficacy and availability of referral and/or diagnostic testing. The largest barriers to uniform medical practice are lack of evidence, conflicting/ambiguous results, and incomplete dissemination of existing information (Eddy 1984, Eddy 1990, Eddy 2007). Primary care has been likened to a jazz performance where some basic structures and heuristics are in place but--owing to uncertainty--practitioners must improvise to a large degree (Miller *et al.* 2001). The next section addresses clinical uncertainty then turns to the contextual factors that play a substantial role in determining the amount, timing, and mixture of diagnostic tests (including imaging) performed in primary care outpatient settings.

Clinical Uncertainty

Any attempt to analyze imaging utilization must first deal with clinical uncertainty. Economic theory provides valuable insight into understanding the role of uncertainty in clinical decision making and makes a clear distinction between uncertainty and risk. Given complete estimates about relevant risks of disease and treatment, patients and doctors could ‘calculate’ an appropriate strategy to maximize expected clinical benefit (Cohen 1996, Wu 1996, Eeckhoudt 1996). Uncertainty, on the other hand, can be defined as ambiguity or imprecision attending risk estimates that renders calculation of expected benefit difficult or impossible. Uncertainty can attach to questions about what (if any) disease the patient has and/or the relative benefit of available treatments for the

disease(s) (Eddy 1984, Wennberg *et al.* 1982). Given uncertainty about disease probability and optimal treatments, practice variations can occur for several reasons (e.g., physician practice style, patient attitudes, organizational influences) and these lead to observed variations in health service utilization. Thus, the evidence-based medicine (EBM) movement is predicated on helping physicians and patients to use medical research findings to reduce their collective uncertainty wherever possible. The hope is to decrease knowledge gaps and information asymmetry about health care using various decision support tools to allow patients and doctors acting as their agents to make better and more consistent clinical decisions.

Assuming that there is a scientific basis for allopathic medicine (i.e., EBM) implies that many diagnoses can be established and optimal treatments known with at least some certainty. Therefore, some of the variation in clinical activity must be due to random differences in disease incidence and prevalence in the population of interest. Routine--evidence based--care of these patients would be expected utilize different amounts and mixtures of services to meet disparate clinical needs. The amount of variation in health service utilization that cannot be explained and justified on the basis of existing and incident disease burden is considered to be unwarranted (Wennberg 2004a, Wennberg 2004b). The method by which expected and unwarranted variation in utilization are parsed from each other involves risk adjustment which seeks to account for the effect of known disease burden on utilization.

Risk Adjustment

Case-mix and risk adjustment methods have been developed primarily to support provider profiling of utilization or patient outcomes (Chang and McCracken 1996, Greene *et al.* 1996, Salem-Schatz *et al.* 1994, Tucker *et al.* 1996, Welch *et al.* 1994).

Consider, for example, the subject of this dissertation: utilization of non-screening diagnostic imaging tests by primary care providers. A payer may calculate rates of utilization and seek to remediate doctors that are high end outliers in order to decrease overall costs. If raw rates of images per patient year are used in this way, targeted providers will--often correctly--object that the reason for their high use is that the patients they see are sicker. This situation clearly calls for case mix adjustment to be 'fair' to targeted doctors, hold patients harmless, and for scientific rigor.

In addition to patient age and gender, outpatient risk adjustment should include variables that capture relevant medical conditions and events. The prevailing method for obtaining such clinical variables is to convert administrative data (claims and/or prescriptions) into problem type and severity categories. The best known example in the U.S., is Ambulatory Care Groups (ACG), a proprietary method developed at Johns Hopkins. Outside the U.S., the most popular method for categorizing primary care case mix is the International Classification of Primary Care (ICPC). Davis used ICPC for case mix adjustment when studying provider variation in primary care practice activity (prescriptions and tests) in Australia (Davis *et al.* 2000a, Davis *et al.* 2002). However, there is little published literature that deals specifically with case mix or risk adjustment for outpatient primary care imaging utilization. One relevant paper deals with 'ambulatory test' utilization (including chest x-ray) by general internists for patients with hypertension (Epstein and McNeil 1985b). Another looked at 'diagnostic services', which included radiology, among generalist and specialists caring for Medicaid patients in community practice (Eisenberg and Nicklin 1981). Both used complex case mix/risk adjustment schemes to account for differences in patient demographics and illness

burden. There is ample literature about case mix (or risk) adjustment in support of performance measurement, resource use comparison, and practice profiling in primary care settings. However, imaging is not the specific focus of these more general methods descriptions (Chang and McCracken 1996, Greene *et al.* 1996, Salem-Schatz *et al.* 1994, Tucker *et al.* 1996, Majeed *et al.* 2001b, Majeed *et al.* 2001a).

Appropriateness and Supplier-Induced Demand

There are several competing theories regarding the causes of medical practice variation that remains after factoring out underlying patient disease burden (i.e., case mix adjustment). An early explanatory model for clinical activity differences depended on categorizing health care utilization events by their 'appropriateness'(Wennberg *et al.* 1982, Wennberg 1987b). The attractiveness of being able to correlate high-intensity utilization with inappropriate services was enhanced by the implication that expenditures could be reduced by targeting high use regions, practices or providers. However, despite considerable effort and expense, no convincing evidence has emerged to show any direct connection between the rate of various types of health service utilization and expert consensus ratings about the appropriateness of care in both large and small regions (Chassin *et al.* 1987, Leape *et al.* 1990, Casparie 1996, Restuccia *et al.* 1996). When focus is narrowed somewhat to primary care providers and their rates of specialist referrals, still no relationship with appropriateness has been demonstrated (Fertig *et al.* 1993, Knottnerus *et al.* 1990).

Another theory, termed Supplier-induced Demand posits that practice variations stem from differences in financial benefits to the supplier (i.e., physician). However, Reinhardt is persuasive in arguing that the theory of supplier induced demand for

inappropriate services fails empirical tests and that a more nuanced explanation is preferable (Reinhardt 1999). He argues for a preferred practice style that:

> would be an amalgam of (1) what the physician has been taught to view as best medical practice in medical school and during residency training, (2) his or her subsequent refinement of the received doctrine on the basis of more recent literature and continuing medical education, and (3) an adaptation to the dominant professional norms in a given locality. (Reinhardt 1999)

That being said, Reinhardt does not deny that financial incentives may influence the "central tendency" of practice patterns over time (Reinhardt 1999). Davis has studied primary care doctors in New Zealand and concludes that a 'supply hypothesis' is not useful and that individual physician's patterns of test ordering and referral form a 'practice style' that persists over time (Davis *et al.* 2000b, Davis *et al.* 2000a, Davis *et al.* 2002). The practice style theory has gained considerable traction and is commonly cited to explain a broad range of variations in medical care delivery (Folland and Stano 1989, Grytten and Sorensen 2003, Welch *et al.* 1993, Wennberg *et al.* 1997, Sirovich *et al.* 2008).

There are scholars of practice variation who argue that the simplistic picture of clinical uncertainty allowing individual practice styles to emerge is theoretically incomplete and does not explain geographic variations because purely personal differences between providers should average out (Stano 1991, Stano 1993, O'Neill and Kuder 2005). However, in some settings, variation in imaging utilization rates persists even when there is clinical certainty. For example, there is strong consensus about breast cancer screening intervals for women over 49 years old and considerable controversy about younger women. In spite of this, small area variations in mammography rates in Ontario are similar across patient age groups (Goel *et al.* 1997).

Westert believes that ‘practice style’ is, at best, shorthand for a cluster of influences acting at the patient, community, and physician level to affect decision making (Westert and Groenewegen 1999). In this more granular and complex model, individual provider opportunities, incentives, and influences combine with shared clinical standards in the group and local medical community to shape their practice. These patterns can be broadly described as ranging between conservative and elaborate which translate into lower and higher use rates for diagnostic tests respectively.

Summary

This dissertation will articulate and quantify factors driving outpatient primary care diagnostic imaging utilization. Further, the study will examine the relative contributions of case mix/clinical need versus contextual/practice style factors to the variation in imaging utilization. Most of the literature about primary care practice variation in test ordering and referral relies on an assumption--sometimes unstated--of proper risk adjustment. To the extent that health service use variations are explained by differences in patient demographics and clinical variables, they should be of less interest to health services researchers (Diehr *et al.* 1990, Cain and Diehr 1992) than to epidemiologists. This epidemiologic/contextual distinction is crucial, even in primary care settings. However, there are relatively few empirical estimates of the relative contributions of these to overall variance. Grytten studied use of diagnostic tests among Norwegian primary care providers (Grytten and Sorensen 2003). The reasons for each visit as well as patient age and gender were used for model-based risk adjustment (clinical need) of diagnostic test expenditures on a per visit basis. The remaining variation ranged between 47-66%, was attributed to practice style, and seemed to be a ‘sticky’ attribute that followed individual providers who changed practice locations during the study.

Other authors have looked at ambulatory primary care practices and estimated the variation in expenditure per patient from clinical need (case mix adjustment) to be about 60% (Phelps *et al.* 1994, Davis *et al.* 2000b).

The next chapter summarizes the research on drivers of resource utilization (emphasizing diagnostic tests and referrals) by primary care providers, then presents a conceptual framework for analyzing imaging utilization in particular. Factors related to clinical need will be addressed first, followed by consideration of the various contextual factors influencing imaging utilization including patient, provider, and practice factors.

CHAPTER 4
CONCEPTUAL FRAMEWORK

This dissertation examines the utilization of outpatient diagnostic imaging by primary care doctors caring for adult patients. Because the study will not consider imaging that occurs as part of disease screening programs, it does not include imaging tests that might be scheduled and performed 'routinely' without an explicit decision by the doctor. Specifically, this study examines diagnostic imaging tests ordered by a patient's primary care doctor to address clinical issues raised either during a visit, a phone call (or email), or another health system encounter (ER visit, hospitalization). This chapter presents a conceptual model to account for relevant driving and modifying factors influencing the amount of such--non-screening--imaging performed by order of the primary care doctor on a patient.

As articulated in the literature review, drivers and modifying factors of primary care resource utilization can be divided into two major classes. The first one is clinical need. For a given patient, clinical need arises from existing or developing signs, symptoms, trauma, and illness. The second class of contextual factors can be further grouped by attribution to the agents or organizations involved. Specifically, most contextual factors will accrue to either the patient or the doctor. Other general groupings for contextual factors include the clinic or practice in which the doctor sees the patient and any larger provider organization (e.g., academic faculty practice, HMO, PPO, and etc). Another category to be considered (at least in the U.S.) relates to the payer or insurer. Other factors relate to the structure and process of outpatient imaging facilities and methods by which imaging tests are ordered. Finally, both patient and doctor live and work in communities that can be represented at various levels of aggregation (e.g., city, county,

state) and the phenomenon of geographic variation implies that at least some factors operate at the community level.

This chapter discusses the relevant driving and modifying factors of outpatient diagnostic imaging utilization in primary care as defined above. For each individual factor, expectations of the direction and strength of effect it should have on imaging utilization are justified by theory and informed by existing empiric literature. The final section summarizes and diagrams the general relationships among clinical need, context, and imaging utilization.

Clinical Need

This set of factors is perhaps the easiest to state and comprehend and yet is the hardest to measure and model. In essence, imaging tests are done to address clinical uncertainty that arises about an existing condition or a clinical event. Existing conditions are diagnoses made by a physician and therefore known to them. Uncertainty about an existing diagnosis relates either to current stage/status of the disease or treatment choice/response for that condition. Clinical events are defined by development or worsening of a sign, symptom, or abnormal test results. The uncertainty engendered by such clinical events comes from the doctor's need to determine if a new diagnosis needs to be made or if an existing disease is responsible. In either case, the function of the diagnostic imaging test is to better identify or exclude treatable disease to guide therapeutic decisions. The only other clinical event not directly accounted for in the preceding explication is trauma, for which imaging is often performed to assess severity and type of injury. Severe trauma is treated in acute care hospitals and associated imaging tests would not be counted as primary care outpatient utilization. On later ambulatory care visits, the 'post traumatic' state will be identified as an existing

condition for this model and outpatient imaging tests might be done to address residual problems.

As described previously, under ideal conditions of evidence-based practice, the mixture and amount of imaging that a patient received in a year would depend only on their existing diagnoses, disease status, and clinical events. However, the current state of medical knowledge and the consistency with which it is applied in actual practice is such that doctor's decisions about intensity and mixture of diagnostic imaging are quite variable even when faced with identical clinical scenarios. This variability in diagnostic decision-making 'styles' among physicians is enabled by systematic uncertainty about what test(s)--if any--are suitable for various clinical scenarios. However, uncertainty or ignorance about the appropriateness (expected clinical benefit) of imaging is necessary but not sufficient for variability in utilization to occur. The contextual factors discussed below, can influence utilization in the face of uncertainty or ignorance about the optimal diagnostic strategy.

For the most part, clinical need variables operate exclusively at the level of individual patients. The only exceptions to this might be in disease screening programs, communicable illness, or environmental health issues. However, this study excludes imaging related to disease screening programs. Furthermore, in the setting for this study (primary care rendered in a large Northeastern metropolitan area from July 2007 through June 2009) no unusual communicable disease epidemics or environmental health issues occurred which might distort the assumption about clinical needs working purely at patient level.

Context: Patient

Basic patient demographics such as gender and age strongly correlate with the amount and type of clinical need based on their complex relationships with many diseases and health states. These two variables are often included in case-mix or risk adjustment models. They serve as proxies for an individual's propensity to develop health conditions and biological responses to disease, testing and treatment. However, these same demographic factors also exert social and psychological effects on the patient's likelihood to seek or accept diagnostic testing under various scenarios. For example, it is conceivable that men and women of similar age might choose differently in some systematic way about diagnostic testing for the same clinical scenario based on level of anxiety related to attitudes about risk and uncertainty.

There is some empiric evidence about patient preferences as related to clinical resource use in general as revealed by surveys, interviews, and responses to hypothetical scenarios. Anthony et. al., found that elderly Medicare beneficiaries expressed substantial differences in their preferences for seeing a doctor right away, having tests, and for specialist care (Anthony *et al.* 2009). When individual Medicare utilization was modeled with these preferences as predictors (along with demographic control variables), those who preferred care right away and from specialists had higher overall healthcare utilization rates. However, at larger levels of aggregation (regional variation) differences in patient preference were uniformly distributed and did not explain variations in cost.

Socioeconomic status (SES) factors such as level of education, income, and ethnicity affect individual patient tendencies to seek care and comply with provider recommendations. Empiric evidence for differences in healthcare utilization and patient

outcomes abounds in the ‘disparities’ literature, although separating ethnicity from economic factors as causes of such disparities is a matter of considerable debate. Confounded with patient’s preferences based on socio-cultural characteristics are the doctor’s own biases and attitudes. Perhaps the best known (and controversial) example of this is Schulman’s survey about recommendation for cardiac catheterization (Schulman *et al.* 1999). He found that, in hypothetical clinical scenarios, patient race and gender influenced physician’s tendency to recommend diagnostic work up for otherwise identical presentations of chest pain. In one intriguing study of managed care claims, Franks found that case mix adjusted use of diagnostic testing was actually higher in patients from lower SES zip codes (Franks and Fiscella 2002). He hypothesized that doctors tend to order more tests when they perceive that patients cannot articulate their histories and current symptoms. However, in general, patients in lower SES and minority ethnicities (in the U.S.) tend to receive lower levels of diagnostic testing in acute and sub-acute care settings (Goldstein *et al.* 2006, Isaacs *et al.* 2004, Pezzin *et al.* 2007, Quintana *et al.* 1997).

Context: Physician

With provider factors, it is important to remember that we are limiting the scope of discussion to outpatient adult primary care. Primary care doctors are trained in several distinct ways in the U.S., and they may choose different types of patients to see. Doctors trained in family medicine often see children and pregnant women in addition to adult patients. On the other hand, geriatrics-trained physicians tend to take care of elderly patients. Many primary care internists have additional training after their 3 years of internal medicine. For example, doctors with some endocrinology training after their internal medicine residency might skew their practice towards adult diabetic patients

even though they are not rendering ‘specialty’ care. However, assuming that patient-level case-mix variables are accounted for, we focus on how factors like experience, gender, training and specialization might affect a doctor’s tendency to order diagnostic imaging in similar clinical scenarios arising in adult primary care.

Training and specialization have direct effects on test-ordering behavior in primary care. Although, all U.S. physicians complete at least a four-year course of study leading to an M.D. degree, specific courses about imaging use are rarely offered or required in M.D. curricula, with the majority of training and experience about radiology gained during residency. Since post-graduate medical education is conducted in ‘apprenticeship’ models, the relative intensity of imaging utilization at the training institution strongly influences subsequent decision making about imaging during practice. Small area variations may be relevant here because in many health care referral regions, the academic medical centers account for and may influence much of the measured utilization. Even within geographic regions, residency training occurs in different types of institutions and community settings. There is wide disparity in these with a spectrum ranging from residencies conducted in small non-affiliated rural community health centers to the classic tertiary care safety net academic health center owned by a medical school operating in a large city. The hypothesis is that doctors trained in high use regions and at large academic centers will tend to order diagnostic imaging more frequently than those trained in smaller centers and low use regions (Chassin 1993, Eisenberg 1986a, Epstein and McNeil 1985a, Folland and Stano 1990, Grytten and Sorensen 2003, Landon *et al.* 2001, O'Neill and Kuder 2005).

Experience is partly defined by length of time in practice and is also directly related to when training took place; the two factors may be very hard to separate. Timing of training is particularly relevant to imaging utilization because the technology has advanced and evolved rapidly and consistently over the past 2-3 decades. The expectation is that a recent graduate who routinely worked with advanced imaging techniques in training will be more likely to order MR and CT in subsequent practice than a doctor trained prior to their diffusion who may not be aware of what is available. Length and extent of experience itself affects diagnostic decision making with doctors having greater experience tending to order less imaging tests in identical scenarios (Bugter-Maessen *et al.* 1996, Childs and Hunter 1972, Couchman *et al.* 2004, Couchman *et al.* 2005, Eisenberg and Nicklin 1981, Sood *et al.* 2007, Whiting *et al.* 2007, Williams *et al.* 1982).

Physician gender also may have an effect on use of imaging independent of experience and training, although, with the recent and substantial increase in the fraction of women trained in and practicing medicine in the U.S., it may be difficult to measure since women physicians tend to be younger and trained later. Empiric data about physician gender has been mixed with studies showing both increased and decreased tendency to order imaging and other diagnostics tests between male and female doctors in primary care (Britt *et al.* 1996, Rosen *et al.* 1997, Sood *et al.* 2007).

Physician workload may influence tendency to order imaging tests in several ways; both in the long run (months and years) and the short run (during the course of a day or week). This effect is mediated through each physician's perception of time pressure overall and during a particular visit. For example, an initial visit to a primary

care physician for headache occurring at the end of a busy day in clinic might be more likely to include an imaging test than otherwise. Such a doctor, caught between time constraint and fear about missing a diagnosis might order an MRI of the head instead of spending an extra 15 minutes doing a detailed neurological exam. On the other hand, the same doctor may choose to refer the patient to a specialist (neurologist in the case of headache) rather than order any imaging if they believe that strategy would move the patient towards a temporary disposition more rapidly. There is very little empiric evidence about this particular factor to guide us in determining the sign of a possible correlation between practice workload and imaging intensity.

Economic factors may influence imaging test ordering (Reinhardt 1999). Aside from pure income maximization, physicians may seek to increase their personal utility in other ways. Eisenberg refers to this as 'physician as self-fulfilling practitioner' (Eisenberg 1986a, Eisenberg 1985). Also, in acting as the patient's agent, physicians may take the patient's financial situation into account when making decisions on their behalf (Eisenberg 1986b). However, imaging utilization also may be influenced by physicians seeking to directly increase their income by ordering and performing imaging tests. In such cases, the reimbursement for the imaging test itself is paid to the ordering physician through several pathways including ownership stake in the imaging equipment (technical fee) and/or interpretation of the examination (professional fee). Called 'self-referral', there is a large body of literature about its practice and ramifications. Hillman and others make a strong empiric and economic case that such direct financial incentives have powerful positive effects on imaging utilization volume and charges (Hillman *et al.* 1990, Hillman *et al.* 1992, Hillman 2004, Gazelle *et al.*

2007). Although the Starke laws have been in effect since the 1990s and have been renewed and revised at least once, recent evidence from California shows that there is still substantial self-referral of advanced diagnostic imaging and that various mechanisms other than direct ownership of equipment allow this arrangement to continue under current statute (Mitchell 2007).

Context: Malpractice

Malpractice deserves special mention because it operates at the individual provider level in addition to the practice and community levels. The term 'defensive medicine' is often used to describe the phenomenon of doctors' increased ordering of imaging--and other diagnostic--tests based on fear of being sued for failure to diagnose (Kessler and McClellan 2002, Sood *et al.* 2007, DeKay and Asch 1998). Personal experience with being sued for malpractice can have profound effects on an individual doctor's psychology and practice pattern that may persist for years or decades. By all accounts, it is an extremely negative and unsettling event that induces a strong desire to avoid repeat occurrences (Hermer and Brody 2010). Thus, if a doctor is sued for missing a diagnosis, the expectation is that they will alter future practice toward more diagnostic testing. This behavior will likely not be limited to the scenario leading to the suit, but generalized across patients of different clinical classes and types of diagnostic tests (including imaging). Even if a doctor is sued for malpractice unrelated to a diagnostic error, he or she may tend to general defensiveness which may lead to greater use of diagnostic imaging at lower levels of uncertainty than before the event.

Dekay and Asch wrote a seminal paper using expected utility theory combined with decision analytic modeling to show the causes and consequences of malpractice liability on diagnostic testing (DeKay and Asch 1998). They proved that consideration of

liability by physicians faced with a classic observe, treat, test choice set must widen the zone (over disease probabilities) in which testing is the preferred strategy. They also prove that there is an obligate utility loss to patients incurred by this extra testing. They also assert that physicians substantially overestimate the 'protection' afforded them by doing more testing. In retrospect, physicians generally overestimate their ability to have made the correct diagnosis in advance. Thus, hindsight and regret bias combine with unrealistic expectations about the efficacy of imaging and result in a near magical belief that the right test would have 'saved the day' (DeKay and Asch 1998).

On a state by state basis, medico-legal 'climate' varies considerably based partly on the statutory and precedent-based status of malpractice and tort laws. Baiker, Fisher, and Chandra's paper examining trends in Medicare costs and malpractice burden in the U.S. over the 1990s used states as the unit of measure (Baicker *et al.* 2007). They showed that imaging cost increases were significantly correlated with trends in malpractice premium and payouts. Across the 50 states, a 10% increase in malpractice premiums/payouts resulted in about two percent increase in physician services costs. They estimated that the observed 10 year increase in malpractice of 60% resulted in more than 15B extra in spending with imaging being the largest contributor by far (Baicker *et al.* 2007). It should be noted that there are diverging views with more recent papers questioning the empiric basis for a large effect by defensive medicine and suggesting that even comprehensive tort reform might not have much actual effect on health costs (Hermer and Brody 2010, Sloan and Shadle 2009).

The theory of how liability concerns increase diagnostic testing at all levels of aggregation rests on the assumption that providers perceive themselves to be at risk for

malpractice action even if they have personally not been sued before (Fenn *et al.* 2007, Kessler *et al.* 2006). Clinicians are rather bad at assessing their own liability risk and tend to overestimate personal probability of being sued (Holtgrave *et al.* 1991, Lawthers *et al.* 1992, Kessler and McClellan 2002, DeKay and Asch 1998). This study will examine a large primary care practice confined to a single institution. Therefore, the local and state malpractice 'climate' is constant though the 'free floating' fear of medical liability might vary by practice. In the current study, the only variable available to directly probe the effect of 'defensive medicine' on imaging utilization is each physician's history of being sued or not during the preceding decade.

Context: Practice Organization

After limiting consideration to adult outpatient primary care in the U.S., there are several types of practice setting and organizational dimensions to be considered. Perhaps the most relevant is the employment arrangement for the physicians. Health care organizations structured as staff models, where doctors are salaried (e.g., traditional HMO, academic health centers, military, and VA) may have different patterns of diagnostic imaging utilization based on individual incentives and medical management initiatives than private practice and fee-for-service settings (Epstein and McNeil 1985a, Kravitz and Greenfield 1995). Even if we exclude consideration of direct financial benefit from self-referral of imaging, independent practitioners and groups are generally less constrained in their ability to order radiology tests.

Aside from employment structures and compensation arrangements, primary care physician practices differ in the extent and manner in which peer pressure is exerted. For example, in a small private practice primary care group there may be minimal (if any) formal influence on actual practice styles among members, including diagnostic

radiology utilization. At the other extreme, in some staff model practices, leadership may routinely profile imaging utilization at the provider level and seek to control it with direct incentives or remedial measures (Neilson *et al.* 2004, Axt-Adam *et al.* 1993, Solomon *et al.* 1998).

Context: Payer and Prices

In the U.S., a patient's insurance status has a profound impact on access to primary care services and may influence the frequency of outpatient visits. In this conceptual model, non screening imaging tests are ordered to address issues identified during a patient-physician encounter, thus a lower frequency of encounters provides fewer opportunities for imaging to be ordered. Additionally, greater financial burden (self-pay or high co-pay/deductibles) associated with imaging tests will reduce a patient's tendency to agree to and/or undergo expensive imaging tests, even if ordered. Physicians may be aware of a patient's financial or insurance status and in their role as financial agents, might choose to forgo imaging tests depending on costs (Mort *et al.* 1996, Shen *et al.* 2004, Pham *et al.* 2007). There is a substantial literature concerning awareness of diagnostic test costs. A recent systematic review by Allen concluded that most doctors have a very limited understanding of diagnostic and non-drug therapeutic costs (Allan and Lexchin 2008). Sood's more focused review of literature about multiple contextual factors in test ordering tendencies found that cost awareness (among both doctors and patients) was relevant (Sood *et al.* 2007). In general, when price information is made available to clinicians, they tend to reduce their likelihood to order diagnostic tests (Hoey *et al.* 1982, Cummings *et al.* 1982, Long *et al.* 1983). Bates reported a 5% decrease in clinical laboratory test charges during inpatient episodes

after price information was routinely displayed during electronic order entry (Bates *et al.* 1997).

Increasingly, various payers (including Medicare), being aware of the rising costs of outpatient imaging tests, have begun to employ cost-containment measures specifically related to imaging (Government Accountability Office (U.S.GAO) 2008). One strategy is to profile individual physicians with respect to imaging (and other resource) utilization and place them into various 'tiers' that give preference to 'efficient' providers in various ways. An emerging trend is for payers to contract with one of several imaging benefits management entities (e.g., National Imaging Associates, CareCore National, and others). These companies serve as 'gatekeepers' for outpatient diagnostic imaging by requiring providers and/or patients to obtain pre-authorization on a case by case basis before tests are scheduled (Otero *et al.* 2006, Brant-Zawadzki 1994, Bernardy *et al.* 2009). In addition to simple barrier effects mediated by call center and other administrative delays, requests for imaging tests may be denied based on proprietary medical necessity or 'appropriateness' rules. Such arrangements can have considerable impact on the likelihood that a given patient-doctor encounter will result in a scheduled and completed diagnostic imaging test (Blachar *et al.* 2006, Levy *et al.* 2006, Smulowitz *et al.* 2009).

Context: Access to Imaging

A final category of contextual factor relates to the facilities and processes that underlie how diagnostic imaging examinations are ordered, authorized, scheduled, performed, and interpreted. As described in the Background, the so-called 'radiology round trip' is a complex chain of events that begins with a doctor-patient interaction of some kind that raises a clinical question that might be answered by imaging. In the case

of a completed examination, the ‘round trip’ generally ends when an interpretative report about the imaging test gets read by the doctor, acted upon, and relayed to the patient.

The availability of diagnostic imaging facilities in terms of proximity and capacity influence the doctor’s tendency to order and the patient’s ability to obtain examinations, even after omitting any consideration of testing facilities owned or operated by the referring physician (self referral). Nonetheless, ready availability and ease of scheduling for various tests will positively influence decision making about imaging by both doctor and patient. The means by which tests are ordered and scheduled by the doctor, office staff, and the patient can introduce barrier or enabling effects. For example, changing from written or verbal orders to a system that requires doctors to log on to a computer and order the test personally, may exert barrier effects if doctors believe that more of their time and effort is required to assert the order. On the other hand, a robust computerized point of care radiology scheduling system can allow patients to leave the clinic with their radiology appointment in hand and will increase utilization by virtue of convenience.

Patient experience at the diagnostic imaging facility may affect compliance with imaging orders as well as a doctor’s tendency to order in the first place. Long wait times and other negative experiences at the testing facility will become known to the doctor and other patients over time. If doctor and patient are contemplating a diagnostic imaging test, expected difficulty in scheduling and/or long waiting times on the day of examination may be perceived as ‘too much trouble’ and bother. Radiologist training, skill and style will affect how they interpret any given test and this is manifest in the report that gets sent back to the referring doctor. If reports tend to be late in arriving,

raise more questions than they answer, and frequently contain recommendations for further testing, doctors may come to rely less on diagnostic imaging. A relatively new development is how the resulting images are handled and distributed to patients and referring doctors. In modern computerized radiology practices, patients are given a CD with all the images on them which can be brought back to the referring doctor for review. Increasingly, images can be viewed on line by the referring doctor along with reports. Imaging providers offer these and other services to increase their market share.

Summary

The final decision of whether or not to order and undergo an imaging test thus depends on all these factors and the complex interactions among them. As shown in Figure 4-1, the ideal level of utilization is determined by clinical need under conditions of certainty. Adding clinical uncertainty then allows for deviation from the ideal level of utilization, with the variation potentially being positive or negative. The various contextual factors add additional variation in the observed level of utilization that persists even after case-mix and risk adjustment.

For purposes of the empirical analysis, the conceptual framework can be summarized as: Imaging utilization = f[clinical need, patient factors, physician factors, malpractice environment, practice organization, payer, access to imaging]

The next chapter will describe the study setting, data sources, and variables that will be used in the analysis.

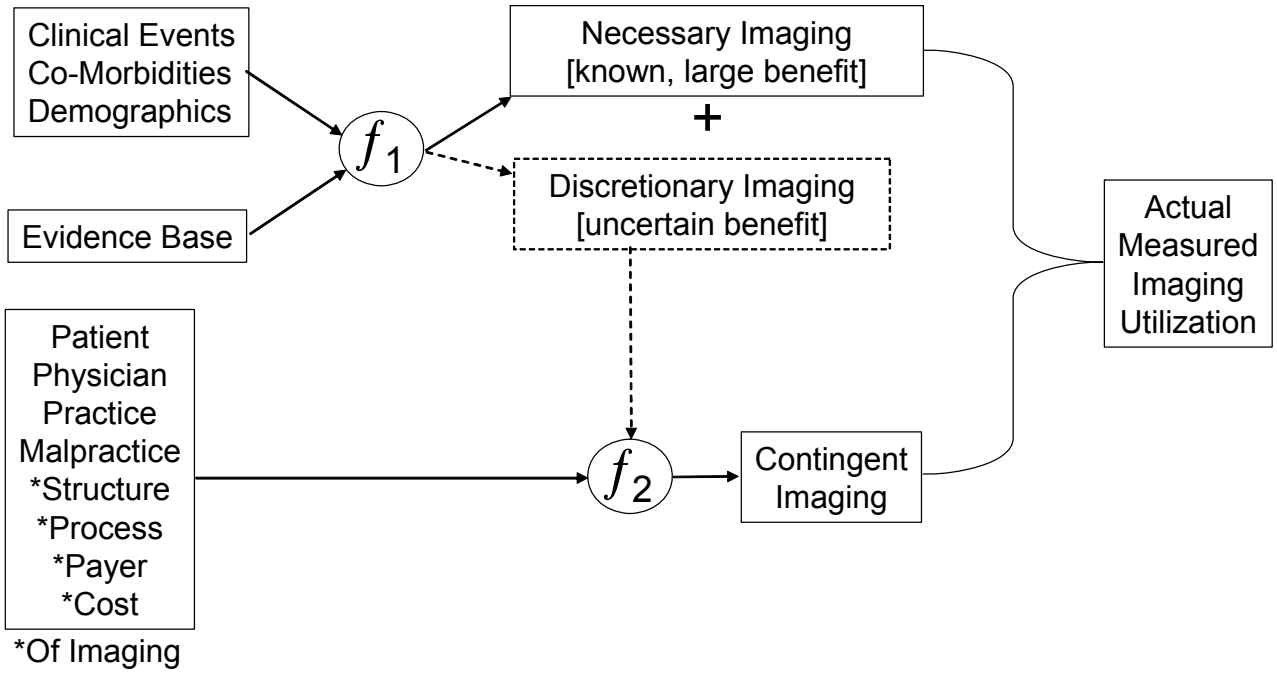

Figure 4-1. Summary diagram of conceptual model for outpatient imaging utilization in primary care.

CHAPTER 5
SETTING, DATA SOURCES, AND VARIABLES

This chapter will describe the institutional setting for this study and then focus on outpatient diagnostic imaging services. The general setting includes a large group of primary care physicians practicing in 15 separate locations and a hospital based radiology department with several service sites.

Patients in the practice who identify with a single attending physician as their primary care provider are tracked as a 'loyalty cohort'. The majority of outpatient imaging performed on these patients occurs at the associated radiology department. Radiology databases provide counts of outpatient diagnostic imaging tests accruing to patients, doctors, and clinics. Other clinical and administrative databases provide information on patient demographics, clinical problems, medical activity (visits, hospitalizations, etc), and physician characteristics.

Settings

Primary Care Practice

This study was conducted at Massachusetts General Hospital (MGH) and the associated Physician Organization (MGPO). In close association with the MGH academic medical center, MGPO is a large multi-specialty faculty group practice with a full complement of adult primary and specialty care. The full time faculty are salaried employees of MGH and the non-radiologists (e.g., primary care providers) have no direct financial incentives relating to volume and/or revenues from imaging or laboratory tests that they order. Malpractice insurance is supplied by the MGPO under a self-insurance pool arrangement and the providers have limited personal liability. However,

malpractice actions are identified at the individual provider level for reporting to licensing authorities and government data-banks.

This work focuses on the primary care portion of the practice and was aided by an entity within the MGPO called the Primary Care Operations Improvement (PCOI) group. Physician leaders and staff members in PCOI conduct analyses of various aspects of the practice for internal quality assessment and improvement as well as for presentation and publication in scholarly settings. In support of these efforts, Dr. Steve Atlas (a primary care physician and health services researcher at MGH) and colleagues have devised and validated a method for identifying a group of primary care patients, doctors, and clinics with stable relationships to each other--the result is termed a loyalty cohort. A cohort is identified by year and comprises a list of patients who are considered to be 'loyal' to a single primary care provider by virtue of their outpatient visits as documented in the electronic medical record over the three years ending in the 'cohort' year. Loyalty assertions for a given patient-doctor pair are calculated probabilistically using five variables derived from visit data (Atlas *et al.* 2006, Lasko *et al.* 2006, Wasiak *et al.* 2008, Atlas *et al.* 2009). These are listed below.

- Waiting fraction: the total number of days waited for appointments with the given physician, divided by the total waited for all physicians combined.
- Visit difference: the total number of visits that a patient has made to the given physician minus the total to all other physicians combined.
- Days since last visit: the number of days since the last visit to the given physician.
- Future difference: the total number of appointments scheduled for future visits with the given physician, minus the total for all other physicians combined.
- Idle ratio: the number of days since the last visit to the given physician, divided by the number of days since the first visit.

A logistic model was developed and validated on subsets using loyalty assertions as stated by the linked primary care physicians (Lasko *et al.* 2006). In comparison with this stated (gold-standard) loyalty status, model predictions were over 95% accurate. When the same technique was applied to all registered primary care patients, the study (loyalty) cohort described below was the result.

Outpatient Radiology

The Radiology Department at MGH provides a full range of imaging services for inpatient, emergency room, and outpatient practices. The main department is located at the MGH campus with several ancillary outpatient sites in the Boston area. The whole department is linked via a robust electronic infrastructure and the radiology informatics group is widely recognized as being among the most advanced and sophisticated in the world. The relevance for this study is two-fold. First, electronic records of all imaging tests are housed in a data warehouse that has been created and maintained with great care and attention to detail. This means that all imaging tests going back to at least 1995 are listed with complete and accurate information about several items relevant to studying outpatient imaging utilization, including the identity of the doctor ordering the test, dates of ordering/completion, modality, body area, and patient status at time of examination (e.g., inpatient, outpatient). In total, almost 100 items of information are stored about each test with the ones just mentioned being most relevant to the present study. Secondly, outpatient radiology ordering by all primary care physicians at MGPO is performed through the same web-based system. This is called Radiology Order Entry (ROE) and it has been in full use for all modalities since 2004. After selecting from a dynamic menu of all outpatient radiology exams, clinicians are required to input

structured information about why the test is being ordered via checkboxes for signs/symptoms/diagnoses supplemented by a free text input field.

A popular feature of the ROE system is a patient scheduling module. This allows an appointment for the imaging test to be made at the point of order (doctor's office) without multiple phone calls, faxes, or other efforts. In addition, starting in 2005, a decision support (DS) component was added to the system that is triggered with all orders for CT, MRI, and nuclear medicine studies. The DS logic displays a 1-9 'utility' score based on the test chosen, patient demographics, and the reasons given for the test. The scores are grouped as follows: 1-3=Red/low, 4-6=Yellow/intermediate, and 7-9=Green/high. A 'red' score does not preclude going ahead and ordering the test but the clinician must provide a reason for doing so prior to proceeding with scheduling. Several papers have been published about the ROE-DS system with the most relevant describing the effect on total outpatient CT, MR, and ultrasound volumes at MGH (Rosenthal *et al.* 2006, Sistrom *et al.* 2009). We found that, after correction for overall practice activity, there were substantial reductions in growth rates, especially for CT scans (Sistrom *et al.* 2009). It should be noted that during the time period (July 2007-June 2009) covered by this study of outpatient imaging utilization in the MGPO primary care practice, the ROE-DS system had been in use for at least two years. Further, no substantial changes were made to the system functionality and only minor alterations were made to the DS scores.

Data Sources

Loyalty Cohort

As described above, the loyalty cohorts are compiled by Dr. Atlas and the PCOI staff every year and the one used in this study is for 2008. This means that MGPO

primary care patients active in 2008 were gathered and their electronic clinic scheduling records going back through 2006 were compiled and analyzed. There were 139,609 unique patients who were candidates for 'loyalty' status. Based on the algorithm, 87,568 were flagged as being loyal to a single primary care provider in the MGPO. There were 804 patients who were loyal to 26 providers with less than 100 loyal patients in their practices. These were excluded leaving 86,764 patients. Of these, 1483 were loyal to four providers who had left the MGPO in late 2008 or the first or second quarters of 2009 and demographic data were not available for 4 of the remaining patients. The analytic sample includes 85,277 patients, loyal to one of 148 primary care physicians who will be characterized below. These physicians practice in one of 15 clinics distributed through the greater Boston area. It should be noted that the clinics sometimes do use residents and medical students. However, the ongoing doctor patient relationship is with the identified primary care physician. In fact, the raison d'etre of the loyalty cohort methodology is to unambiguously identify this relationship.

Patient Details

MGH and the MGPO have established a common Research Patient Data Repository (RPDR), which aggregates numerous sources of information into a single set of databases designed for use in patient-centered clinical epidemiology. These include billing and encounter data for inpatient, emergency department, and outpatient services as well as the contents of inpatient hospital information systems (HIS) and outpatient electronic medical records (EMR). This study was approved by the Institutional Review Board at MGH under an expedited protocol for analysis of existing data. Informed consent was not required and was not obtained.

Physician Details

The MGPO credentialing database was used to obtain relevant information pertaining to the 148 primary care doctors in the study, including gender, birth year, medical school graduation year, and medical school state (or country for foreign medical graduates). By definition, all doctors were licensed in the state of Massachusetts and the publicly available web site for the Massachusetts Board of Medicine was queried to determine whether or not each doctor had been sued for medical malpractice in the past 10 years.

Imaging Utilization

The 85,277 patient medical record numbers were queried against the radiology data warehouse to return all diagnostic imaging tests performed during the study interval (July 1, 2007 through June 30, 2009). The query specifically excluded interventional procedures (e.g., biopsy, drainage, catheter angiography, embolization, and vascular stenting among others). Also specifically excluded were mammograms, as these are almost all related in some way to breast cancer screening. In addition to patient medical record number, modality, body area, place of service (ER, inpatient, outpatient), and the unique provider number of the doctor who ordered the exam was obtained.

The 221,571 diagnostic imaging procedures performed on cohort patients over the two-year study interval cross tabulated by place of service and modality are shown in Table 5-1. The 157,463 outpatient diagnostic imaging procedures are cross tabulated by the class of the ordering doctor and modality in Table 5-2. Finally, the 60,938 diagnostic imaging procedures performed in the outpatient setting and ordered by the

patient's linked (loyal) doctor are cross tabulated by body area imaged and modality in Table 5-3.

The unit of observation for this study is the patient. The outcome variable is constructed by aggregating and summing (by patient) the 60,983 outpatient imaging procedures ordered by the primary care provider to whom the patient was loyal in 2008. The remaining diagnostic imaging procedures were performed while patients were in the emergency room (N=34,345), completed while patients were in the hospital (29,763), or ordered as outpatient by providers other than the patient's loyal doctor (other primary care=9,833, specialists=86,692). These other categories of imaging utilization were also aggregated by patient and summed to produce the other patient-level imaging utilization (independent) variables described below.

Variables

Imaging Utilization (dependent variable)

The main patient level outcome variable is called prv_o_cnt and is the count of the number of outpatient diagnostic imaging tests (CT, MR, NM, PET, X-Ray, US) ordered by the provider to whom the patient was loyal during the study period (July 1, 2007 through June 30, 2009). Univariate statistics for this variable are summarized in Table 5-4.

Patient Characteristics

There were 35,709 men (41.9%) in the cohort whose ages in 2008 ranged from 17-100 with mean=54.5 years and standard deviation of 15.1 years. There were 49,568 women (58.1%) whose ages in 2008 ranged from 17-103 with mean=53.2 years and standard deviation of 16.4 years. Patient race was available for all subjects and is shown in Table 5-5.

Each patient's payer of record (in 2008) was available in the RPDR as obtained from outpatient billing systems. These were initially categorized into 13 levels (Table 5-6).

For modeling purposes, the 13 payer levels were collapsed into 6 levels as follows:

- Aetna into Commercial
- Harvard Pilgrim Healthcare, Neighborhood Health Plan, and Tufts Health Plan into Managed
- Mass Health Net and Medicaid into State
- Free Care and Self Pay into Uninsured

The resulting 6 level insurance payer categories (Table 5-7) will be used for all subsequent analyses.

Clinical Events

The RPDR was used to obtain counts of various clinical events for each patient. These were summed over the period from July 1, 2007 through June 30, 2009. The events counted all occurred at MGH. Hospital activity variables (summarized in Table 5-8) included visits to the emergency room, inpatient hospital stays, ICU days, and inpatient observation stays. Observation (short) stays are a special category of hospitalization where the patient remains in the hospital for less than 24 hours. Observation stays often occur in concert with an emergency room visit and allow for extended nursing care without a formal admission. Observation stays also are used for minor procedures, dialysis, and administration of intravenous medications. Outpatient visit counts for each patient during the study period (July 1, 2007 through June 30, 2009) were obtained from the RPDR and confined to the 15 CPT codes representing outpatient office visits. The professional RVU for each of these CPT codes during 2008

was obtained from the CMS website (see Table 5-9). In addition to counting the number of outpatient visits, the RVU of those visits were also summed to form a separate variable. The visit counts and summed RVU for visits were stratified by type of doctor being visited (prv=provider to whom patient was loyal, pcp=other primary care doctor, spc=specialist). The resulting six variables are summarized in Table 5-10.

Clinical Problems

The EMR systems used by primary care providers to document outpatient care of all patients in the study cohort allow for recording of a 'problem list' for each patient. These problems are encoded in one of two ways depending on the clinic. The coding systems are internal to MGH and crosswalk tables are available to parse the problem codes into broad categories, including diabetes, hypertension, heart failure, coronary artery disease, renal failure, cancer, trauma, obesity, and substance abuse. Active problems not falling into one of these groups were labeled as 'other problem' for purposes of this study. For each of the major categories, a binary variable was constructed with value 'true/yes/1' when the patient had at least one active problem listed in their EMR entries falling into that category. That same variable was assigned with 'false/no/0' when the patient did not have an active problem asserted falling into the category in question. The nine binary clinical problem variables are summarized in Table 5-11.

In the cohort of 85,277 patients, 46,063 (54.0%) had none of the problem categories listed above, 23,265 (27.3%) had 'yes' for a single category, 10,808 (12.7%) had two of the problem categories asserted positively, 3,846 (4.5%) had 'yes' for three categories, 991 (0.30%) had four positive categories, and 304 (0.36%) had 'yes' for five or more.

The EMR systems also included additional problem assertions for many patients that did not fall into one of the categories listed above. Examples include depression, hepatitis, arthritis, and so forth. For each patient, the counts of unmapped problem codes were placed into a variable of other problems (name=oth_prob). Overall summary statistics for the count of other problems included median of 6.0, mean of 7.75, and standard deviation of 6.94. This count of other problems was zero for 5709 patients and of those, 4941 (5.8% of the whole cohort) had none of the problem categories enumerated above (e.g., patient's clinical problem list was empty/null).

Outpatient Prescriptions

The EMR systems also serve prescribing and drug reconciliation functions for the primary care practice. In the cases in which the patients do not have e-prescribing enabled to their pharmacy and/or drug benefits program, orders for outpatient medications are still entered into the EMR and printed prescriptions are given to patients. To enumerate the number of outpatient medications each patient was taking during 2008, the number of 'active' prescriptions was counted starting from the first available entry for any given patient. The queries did not count refills of the same drug and dose as new prescriptions. However, switches within a drug class and/or dose changes were counted as new prescriptions which could result in over-counting and rendering small differences in the discrete number less meaningful than the general amount each patient was taking. Therefore, we stratified the count of active outpatient medications into four categories (variable name=meds_cat) summarized in Table 5-12.

Other Imaging Utilization

Counts of diagnostic imaging tests performed on each patient during the study interval (July 1, 2007 through June 30, 2009) were stratified by place of service (i.e.,

patient status/location at the time the test was performed). These include emergency department, inpatient, and outpatient (Table 5-1). Outpatient exams were further stratified by the category of physician ordering them; primary care physician (other than the patient's own loyal doctor), and specialist (Table 5-2). Note that the counts of outpatient diagnostic imaging tests ordered by the patients linked (loyal) provider is the outcome variable which has already been described above. Summary statistics for the four strata of (non-outcome) imaging utilization are listed in Table 5-13.

Physician Characteristics

The cohort of 85,277 patients were 'loyal to' 148 primary care physicians; 76 (51.3%) women and 72 (48.7%) men. A variable called prov_age_08 was constructed using each doctor's birth year. The ages of the male physicians (in 2008) had minimum=33, maximum=75, mean=49.7, and standard deviation=10.0 years respectively. The female physicians ages (in 2008) had minimum=31, maximum=63, mean=45.4, and standard deviation=8.4 years respectively. The year of medical school graduation was used to construct a variable called prov_exp_08 that quantifies the number of years between medical school graduation and 2008. This (prov_exp_08) had minimum=5, maximum=50, mean=19.6, and standard deviation=9.6 years respectively. Perhaps a better proxy for physician experience would have been years since completion of residency. However, the credentials database had incomplete and inconsistent information in this regard.

The number of (loyal and non-loyal) patients in the cohort linked to each doctor was summed into a variable called prov_pat_count and it had minimum=172, maximum=2394, mean=801.3, standard deviation=396.0, and median=751. This variable (prov_pat_count) was categorized into four levels (variable

name=prov_pat_cat) and will be used as a proxy for how busy the doctor was during the study period (see Table 5-14).

The medical school graduation state/country was used to construct a variable called prov_fmg that was set to ‘yes’ (N=8, 5.4%) when the doctor had graduated from a medical school outside the US (Argentina=1, Canada=2, Croatia=1, England=1, Holland=1, Italy=1, Panama=1). The graduate level degrees held by each provider were used to construct a variable called prov_md_plus that was set to ‘yes’ (N=14) when the doctor had obtained a graduate degree in addition to their M.D. (MPH=6, MSC=1, MSW=1, PHD=6). The assertions about malpractice cases in the last 10 years found on the Massachusetts Board of Medicine web site were used to construct a variable called mp_flag with ‘yes’ (N=7) when the doctor had a record of having been sued and ‘no’ otherwise (N=147).

Site Characteristics

Each of the 15 sites was labeled with a unique number that will serve as identification for subsequent analysis about inter-site variability. The only variable that accrues to the clinics themselves is the number of primary care doctors in active practice at each one. This variable along with the number of patients and primary care doctors assigned to each site serve as proxies for the ‘size’ of the clinics. These are summarized in Table 5-15.

Note that the sum of the number of active doctors at all sites is 168 whereas there were only 148 included in this study (Study Doctors Column). The remaining 20 had less than 100 loyal patients and were not included in the analytic data set.

Patient, Provider, and Clinic Identifiers

For multi-level (hierarchical) modeling it is necessary to uniquely identify the individual primary care doctors being studied so as to be able to maintain the linking between them and their loyal patients. Therefore, the MGH provider identifying numbers were sorted and mapped to the corresponding rank (1, 2, 3, … , 148). Using this mapping, each of the 85,277 patient's linked MGH provider identifier was replaced by a unique (though now anonymous) integer. Since each observation in the final analytic data set represents a single patient, no identifying information need be retained (e.g., Name, Medical Record Number, and etc) and these were all dropped. As described above, the clinics were already identified by (uninformative) integers (1-15).

Data Integrity: Clinical Activity Variables

One way to insure that the queries of outpatient visits, inpatient stays, emergency room visits, and outpatient imaging tests for all patients in the study were complete and consistent is to plot them over time. This was done by counting each event type by month for the whole study cohort and plotting as a time series. The visit counts by month stratified by type of provider rendering the visit are plotted in Figure 5-2.

It is reassuring that the counts are relatively stable and consistent during the study period. This implies that there are no large gaps or duplications in the data. As for any secular patterns, this dissertation will not attempt to describe or explain them.

Similarly, the counts of outpatient diagnostic imaging examinations were enumerated by month and stratified by the type of provider ordering the study are plotted in Figure 5-3.

As with the counts of outpatient visits, there is apparent consistency and stability. Further reassurance comes from the fact that relative decreases in counts during

December 2007 and August 2008 seem to match between outpatient visits and imaging tests. Since the data came from separate and independent administrative sources, there was an actual decrease in outpatient clinical activity during these periods. Visual inspection of the two plots confirms that visits and imaging tests tend to rise and fall together by month.

The other main activity variables relate to hospital events and counts of these by month are plotted in Figure 5-4.

Again the month to month consistency and stability attests to the integrity of the data, which came from two separate databases (one for emergency room and a second for inpatient/observation stays).

Variable Summary

A summary of all patient-level independent variables is provided for reference in Table 5-16. A summary of all the clinic and provider level independent variables is provided for reference in Table 5-17.

Table 5-1. All diagnostic imaging performed on study cohort during two years of study.

Place Of Service	CT	MR	NM	PET	X-Ray	US	Total	Percent
ER	10183	2651	489	8	18608	2406	34345	15.50
Inpatient	4768	1634	944	128	19907	2382	29763	13.43
Outpatient	26981	21999	6129	3709	76654	21991	157463	71.07
Total	41932	26284	7562	3845	115169	26779	221571	100
Percent	18.92	11.86	3.41	1.74	51.98	12.09	100	-

Note: CT=computed tomography, MR=magnetic resonance imaging, NM=nuclear medicine, PET=positron emission tomography, X-Ray=radiography, US=ultrasound. The studies performed in the outpatient setting are further stratified in Table 5-2.

Table 5-2. Outpatient diagnostic imaging performed on study cohort during two years of study.

Who Ordered	CT	MR	NM	PET	X-Ray	US	Total	Percent
Specialist	15295	13170	3684	3294	42143	9106	86692	55.06
Other primary care doctor	1133	820	157	17	6605	1101	9833	6.24
Patient's own loyal doctor	10553	8009	2288	398	27906	11784	60938	38.70
Total	26981	21999	6129	3709	76654	21991	157463	100
Percent	17.13	13.97	3.89	2.36	48.68	13.97	100	-

Note: CT=computed tomography, MR=magnetic resonance imaging, NM=nuclear medicine, PET=positron emission tomography, X-Ray=radiography, US=ultrasound. The studies ordered by the patient's linked (loyal) doctor are further stratified in Table 5-3.

Table 5-3. Outpatient diagnostic imaging ordered by patient's linked (loyal) doctor during two years of study.

Body Area	CT	MR	NM	PET	X-Ray	US	Total	Percent
Abdomen	4005	659	25	89	762	4032	9572	15.71
Cardiac	42	10	2003	0	0	0	2055	3.37
Chest	4637	303	13	96	11084	12	16145	26.49
Extremity	118	1807	0	0	10776	834	13535	22.21
Head/Brain	976	2060	0	4	244	152	3436	5.64
Maxillofacial and/or Neck	325	249	54	95	578	2055	3356	5.51
Pelvis	256	318	0	0	1097	4310	5981	9.81
Spine	193	2602	0	0	3355	0	6150	10.09
Unspecified	1	1	193	114	10	389	708	1.16
Total	10553	8009	2288	398	27906	11784	60938	100
Percent	17.32	13.14	3.75	0.65	45.79	19.34	100	

Note: CT=computed tomography, MR=magnetic resonance imaging, NM=nuclear medicine, PET=positron emission tomography, X-Ray=radiography, US=ultrasound.

Table 5-4. Univariate statistics of the outcome variable (per patient count of outpatient imaging tests ordered by primary care provider).

N	85277
Minimum	0 (N=53,617)
Maximum	15
Mean	0.7146
Standard Deviation	1.2563
Skewness	2.6554
Uncorrected SS	178132
Coeffient of Variation	175.8049
Sum of Observations	60938
Variance	1.5782
Kurtosis	9.8796
Corrected SS	134586
Standard Error of the Mean	0.0043

Note: SS=sum of squares.

Table 5-5. Distribution of patient race.

Race	Frequency	Percent
White	68432	80.25
Black	4278	5.02
Hispanic	5924	6.95
Other	6643	7.79
Total	85277	100

Table 5-6. Distribution of patient's payer categories.

Payer Group	Frequency	Percent
Aetna	1717	2.01
Blue Cross Blue Shield	30762	36.07
Commercial	6282	7.37
Free care	196	0.23
Harvard Pilgrim Healthcare	8574	10.05
Mass Health Net	1890	2.22
Medicaid	3646	4.28
Medicare	19555	22.93
Neighborhood Health Plan	1573	1.84
Other	3057	3.58
Self Pay	1057	1.24
Tufts Health Plan	6968	8.17
Total	85277	100.00

Table 5-7. Patient's payer collapsed into 6 categories.

Payer Group	Frequency	Percent
Blue Cross Blue Shield	30762	36.07
Commercial	7999	9.38
Managed	17115	20.07
Medicare	19555	22.93
Other	3057	3.58
State	5536	6.49
Uninsured	1253	1.47
Total	85277	100.00

Table 5-8. Hospital activity variables (per patient).

Description	Variable Name	Minimum	Maximum	Mean	SD
Total hours in Emergency room	er_hours	0 N=69,485	483	2.71	10.15
Emergency room visits	er_visits	0 N=69,484	47	0.33	1.04
Inpatient admissions	inpt_stays	0 N=74,582	22	0.20	0.67
total days in hospital	Inpt_los_total	0 N=74,697	179	0.92	4.70
Days in intensive care units	inpt_icu_days	0 N=84,401	49	0.05	0.76
Readmitted within two weeks of inpatient discharge	inpt_read_15d	0 N=84,382	10	0.01	0.15
Readmitted within one month of inpatient discharge	inpt_read_31d	0 N=84,062	17	0.02	0.21
Observation (short) stays	obs_stays	0 N=76,145	37	0.13	0.54

Note: The number of patients with zero counts is given under Minimum (where zero).

Table 5-9. CPT codes and relative value units for ambulatory office visits.

CPT Code	Visit Type	Complexity	RVU
99201	new patient	not comprehensive not complex	0.48
99202	new patient	not comprehensive mod complex	0.93
99203	new patient	not comprehensive high complex	1.42
99204	new patient	comprehensive moderate complexity	2.43
99205	new patient	comprehensive high complexity	3.17
99211	established	not comprehensive not complex	0.18
99212	established	not comprehensive mod complex	0.48
99213	established	not comprehensive high complex	0.97
99214	established	comprehensive moderate complexity	1.50
99215	established	comprehensive high complexity	2.11
99241	consultation	not comprehensive not complex	0.64
99242	consultation	not comprehensive mod complex	1.34
99243	consultation	not comprehensive high complex	1.88
99244	consultation	comprehensive moderate complexity	3.02
99245	consultation	comprehensive high complexity	3.77

Table 5-10. Outpatient visit activity variables (per patient).

Description	Variable Name	Minimum	Maximum	Mean	SD
Count of outpatient visits to linked (loyal) provider	prv_visit_count	0 N=10,396	62	3.54	3.39
Summed RVU of visits to linked (loyal) provider	prv_visit_rvu	0 N=10,396	76.22	4.95	4.83
Count of outpatient visits to other primary care doctors	pcp_visit_count	0 N=57,117	75	0.60	1.24
Summed RVU of visits to other primary care doctors	pcp_visit_rvu	0 N=57,117	59.44	0.73	1.57
Count of outpatient visits to specialists	spec_visit_count	0 N=25,771	112	3.60	5.17
Summed RVU of visits to loyal specialists	spec_visit_rvu	0 N=25,771	165.18	4.96	7.28

Note: The number of patients with zero counts or RVU is given under Minimum (where zero).

Table 5-11. Binary clinical problem variables (per patient).

Problem category	Variable Name	Number Yes	Percent Yes
Diabetes	pr_dm	9485	11.12
Hypertension	pr_htn	25219	29.57
Heart failure	pr_chf	867	1.02
Coronary artery disease	pr_cad	3940	4.62
Renal failure	pr_crf	1096	1.29
Cancer	pr_can	9925	11.64
Trauma	pr_trm	1804	2.12
Obesity	pr_obs	8855	10.38
Substance abuse	pr_sub	770	0.90

Table 5-12. Four level categorization of patient active outpatient medications (per patient).

meds_cat	Frequency	Percent
None	4475	5.25
1-5	45536	53.40
6-10	23390	27.43
>10	11876	13.93
Total	85277	100.00

Table 5-13. Summary of other (non-outcome) imaging test utilization variables (per patient).

Description	Variable Name	Minimum	Maximum	Mean	SD
Count of imaging tests ordered during emergency room visits	all_e_cnt	0 N=73,562	96	0.40	1.56
Count of imaging tests ordered during inpatient stays	all_i_cnt	0 N=79,118	118	0.35	2.41
Count of outpatient imaging tests ordered by specialists	spec_o_cnt	0 N=55,105	61	1.02	2.28
Count of outpatient imaging tests ordered by other primary care doctors*	pcp_o_cnt	0 N=78,289	8	0.12	0.45

Note: The number of patients with zero counts is given under Minimum (where zero).
*Not the patient's own linked (loyal) doctor.

Table 5-14. Four level categorization of the number of patients cared for by each provider (panel size).

prov_pat_cat	Frequency	Percent
<500	37	25.00
500-759	36	24.32
750-999	36	24.32
1K+	39	26.35

Table 5-15. Site (clinic) characteristics.

Site (clinic) ID	Active Doctors	% Active Doctors	Study Patients	% Study Patients	Study Doctors	% Study Doctors
1	7	4.17	3451	4.05	5	3.38
2	8	4.76	4233	4.96	8	5.41
3	20	11.90	11925	13.98	16	10.81
4	11	6.55	5058	5.93	9	6.08
5	15	8.93	7220	8.47	13	8.78
6	5	2.98	3269	3.83	4	2.70
7	7	4.17	3308	3.88	6	4.05
8	17	10.12	10381	12.17	17	11.49
9	12	7.14	6378	7.48	12	8.11
10	14	8.33	5728	6.72	14	9.46
11	9	5.36	5097	5.98	7	4.73
12	18	10.71	9109	10.68	16	10.81
13	5	2.98	1135	1.33	5	3.38
14	6	3.57	4794	5.62	6	4.05
15	14	8.33	4191	4.91	10	6.76
Total	168	100	85277	100	148	100

Table 5-16. Description and categorization of 33 patient level independent variables.

Variable Name	Level	Class	Type	Levels	Description
inpt_stays	Patient	Activity, Hospital	Numeric	-	count of inpatient stays
er_hours	Patient	Activity, Hospital	Numeric	-	total hours spend in er
er_visits	Patient	Activity, Hospital	Numeric	-	count of er visits
obs_stays	Patient	Activity, Hospital	Numeric	-	count of observation stays
inpt_read_31d	Patient	Activity, Hospital	Numeric	-	count of readmit within 31 days
inpt_read_15d	Patient	Activity, Hospital	Numeric	-	count of readmit within 15 days
lnpt_los_total	Patient	Activity, Hospital	Numeric	-	total days in hospital
inpt_icu_days	Patient	Activity, Hospital	Numeric	-	total days in icu
all_e_cnt	Patient	Activity, Other Imaging	Numeric	-	count of images done in ER
all_i_cnt	Patient	Activity, Other Imaging	Numeric	-	count of images done as inpatient
spec_o_cnt	Patient	Activity, Other Imaging	Numeric	-	count of outpatient images ordered by specialists
pcp_o_cnt	Patient	Activity, Other Imaging	Numeric	-	count of outpatient images ordered by other primary care doctors
prv_visit_rvu	Patient	Activity, Visits	Numeric	-	sum of rvu of outpatient visits to linked (loyal) primary care physician
prv_visit_count	Patient	Activity, Visits	Numeric	-	count of outpatient visits to linked (loyal) physician
spec_visit_rvu	Patient	Activity, Visits	Numeric	-	sum of rvu of outpatient visits to specialists
spec_visit_count	Patient	Activity, Visits	Numeric	-	count of outpatient visits to specialists
pcp_visit_rvu	Patient	Activity, Visits	Numeric	-	sum of rvu of outpatient visits to covering pcp
pcp_visit_count	Patient	Activity, Visits	Numeric	-	count of outpatient visits to covering pcp
age_08	Patient	Demographics	Numeric	-	Patient’s age in 2008
Race	Patient	Demographics	Categorical	4	patient identified race
Sex	Patient	Demographics	Categorical	2	patient sex
payer_group	Patient	Insurance	Categorical	6	category of patient's payer of record in 2008
meds_cat	Patient	Medications	Categorical	4	active outpatient prescriptions in 2008
pr_cad	Patient	Clinical Problem	Categorical	2	coronary artery disease
pr_can	Patient	Clinical Problem	Categorical	2	cancer
pr_chf	Patient	Clinical Problem	Categorical	2	congestive heart failure
pr_crf	Patient	Clinical Problem	Categorical	2	chronic renal failure
pr_dm	Patient	Clinical Problem	Categorical	2	diabetes
pr_obs	Patient	Clinical Problem	Categorical	2	obesity
pr_htn	Patient	Clinical Problem	Categorical	2	hyptertension
pr_sub	Patient	Clinical Problem	Categorical	2	substance abuse
pr_trm	Patient	Clinical Problem	Categorical	2	trauma
oth_prob	Patient	Clinical Problems	Numeric	-	count of active problems other than those separately listed above

Table 5-17. Description and categorization of provider (8) and clinic level (2) independent variables.

Variable Name	Level	Class	Type	Levels	Description
site_docs	Clinic	Characteristic	Numeric	-	number of doctors actively practicing at the clinic in 2008
site_id	Clinic	Identifier	Categorical	15	anonymous clinic (site) identifier
mp_flag	Provider	Characteristic	Categorical	2	whether provider has been sued in last 10 years
prov_md_plus	Provider	Characteristic	Categorical	2	whether provider has a degree beyond MD
prov_pat_cat	Provider	Characteristic	Categorical	4	number of patient's in provider practice in 2008
prov_sex	Provider	Characteristic	Categorical	2	provider sex
prov_fmg	Provider	Characteristic	Categorical	2	whether provider is foreign medical graduate
prov_age_08	Provider	Characteristic	Numeric	-	age in years of the provider in 2008
prov_exp_08	Provider	Characteristic	Numeric	-	number of years after provider MD graduation in 2008
prov_id	Provider	Identifier	Categorical	148	anonymous provider identifier

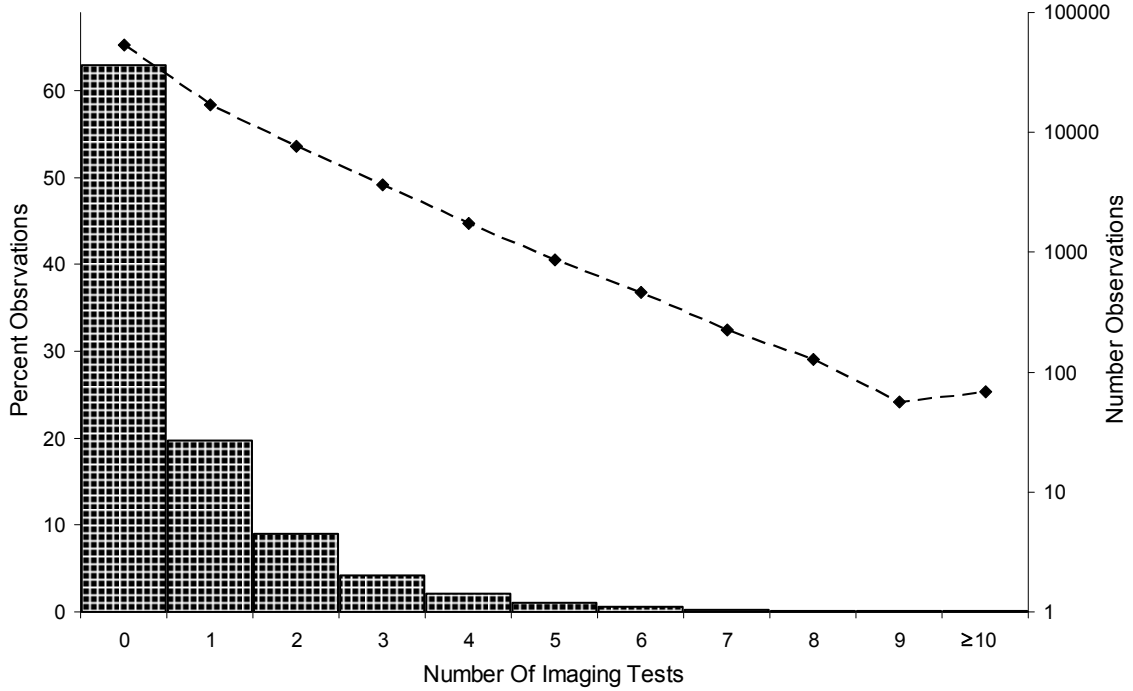

Figure 5-1. Outpatient imaging tests (per patient) ordered by the linked (loyal) primary care provider. Both the percent (left Y axis, bars) and number (right Y logarithmic axis, diamonds and dashes) of observations are shown.

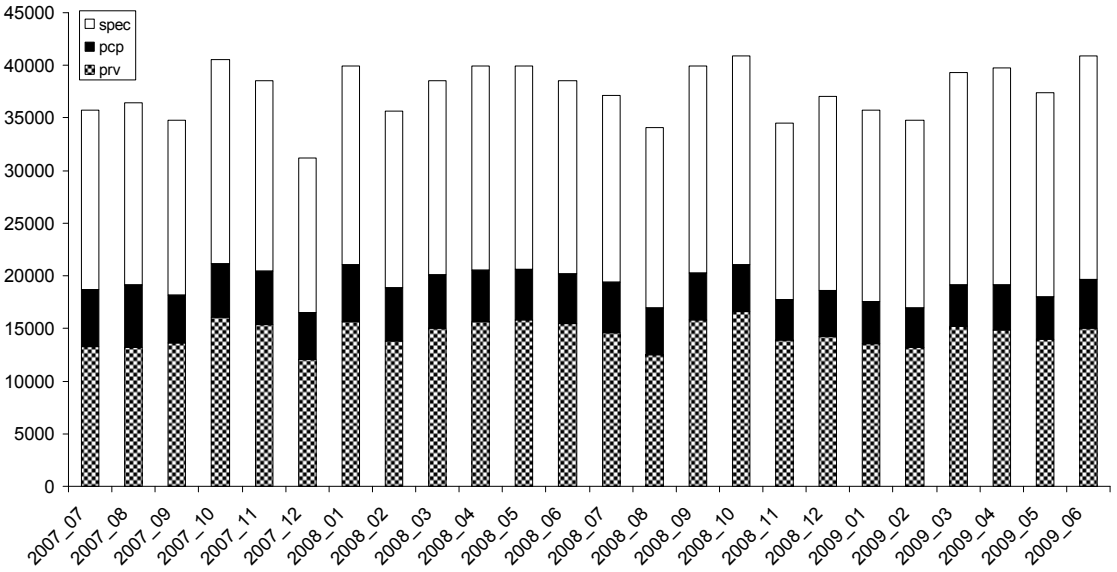

Figure 5-2. Number of outpatient visits by all patients in study cohort (by month) over two years of study. Hatched: visits to patient's linked (loyal) doctor, Black: visits to another (covering) PCP, White: visits to specialists.

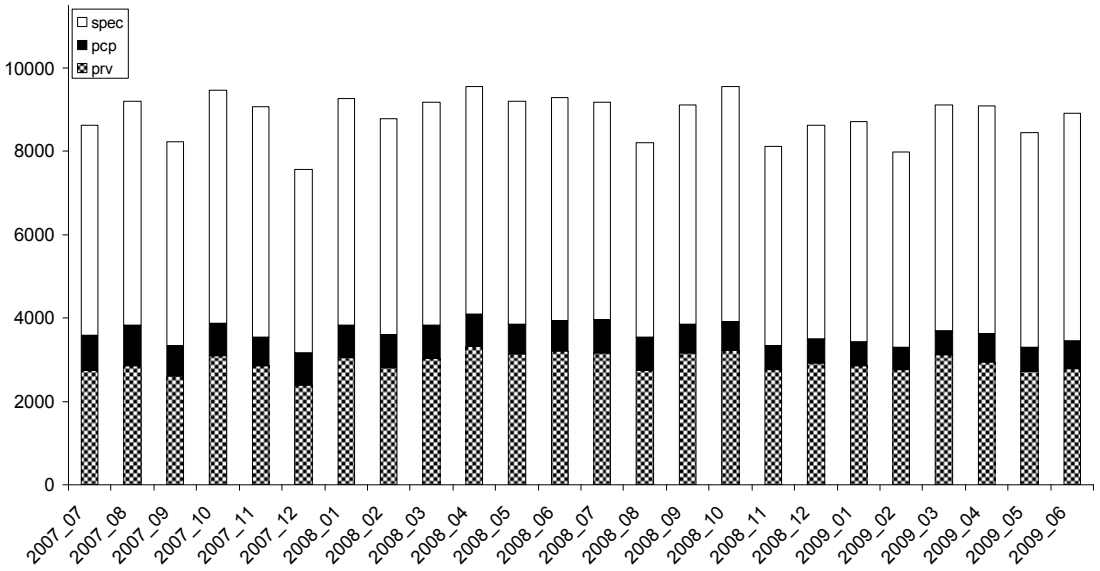

Figure 5-3. Number of outpatient imaging tests performed on all patients in study cohort (by month) over two years of study. Outpatient imaging tests ordered by: Hatched: patient's linked (loyal) doctor, Black: another (covering) PCP, White: specialists.

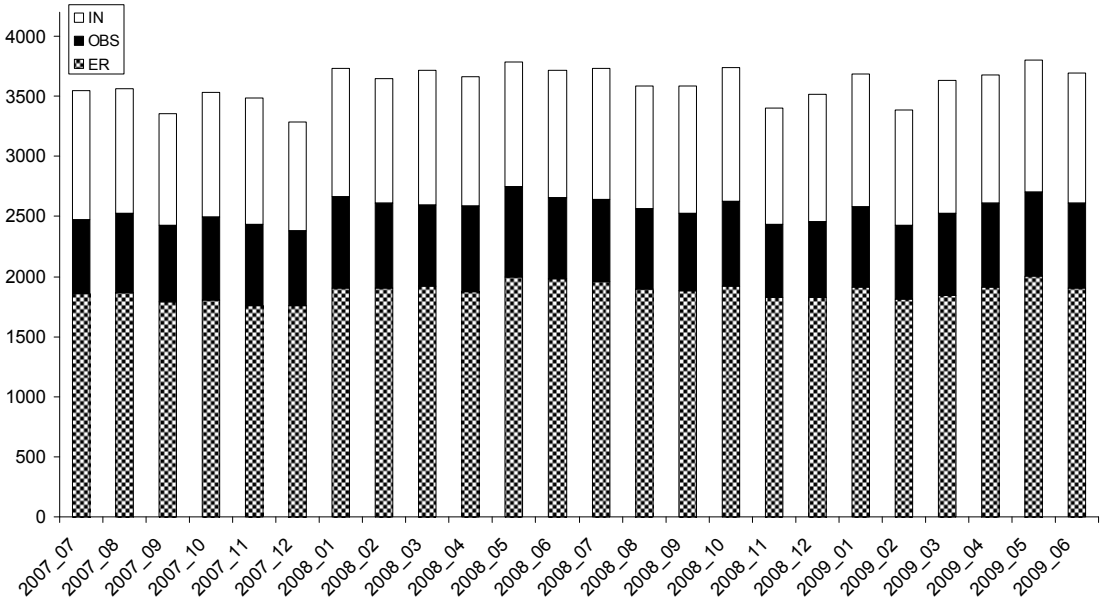

Figure 5-4. Number of hospital encounters for all patients in study cohort over two years. Hatched: emergency room visits, Black: short (observation) stays, White: inpatient discharges.

CHAPTER 6
METHODS

Outcome Variable Distribution

Simple univariate statistics of the main outcome variable (prv_o_cnt = IMG for this chapter) have been described above in Chapter 5 (see Table 5-4 and Figure 5-1). Based on the visual inspection, these appear to be count data with a Poisson distribution. This is a discrete distribution, with a single parameter, Lambda (λ), that expresses the probability of a number of events occurring in a fixed period of time if these events occur with a known average rate (λ) and independently of the time since the last event. Each instance of a Poisson random variable can be expressed as being the result of a single Poisson 'experiment' and may take any positive integer value (Stahl 1969). An important property of the Poisson distribution is that the variance is equal to the mean (λ). The Poisson equation may be written as Equation 6-1.

$$\text{Prob}(n) = (\lambda^n e^{-\lambda}) / n! \qquad (6\text{-}1)$$

In Equation 6-1, e is the base of the natural logarithm, n is the number of occurrences of an event - the probability of which is given by the function, and λ is the positive real number, equal to the expected number of occurrences that occur during the given interval.

A more theoretically appealing and possibly informative way to describe the distribution of imaging counts is that of a two stage process with the first stage determining the occurrence of any imaging from the whole data set of 85,277 patients and the second stage determining the count of imaging tests for patients that had at least one (IMG > 0, N=31,660) imaging test during the study period. This can be modeled with a logistic regression on all observations followed by Poisson regression

on the non-zero observations. For purposes of estimating patient, provider, and clinic effects; the two step (logistic>Poisson) approach was taken.

A final method to model the distribution of outpatient imaging counts combines the logistic portion (any use) with the Poisson assumption for intensity (non zero use) into a single distribution; zero inflated Poisson (ZIP). The ZIP distribution is almost the same as the standard Poisson when n>0 but has a second portion for n=0 and may be written as Equations 6-2 and 6-3.

$$\text{Prob}(n) = \phi + (1-\phi)\, e^{-\lambda} \qquad \text{for } n = 0 \tag{6-2}$$

$$\text{Prob}(n) = (1-\phi)\, (\lambda^{n}\, e^{-\lambda}) / n! \qquad \text{for } n = 1, 2, \ldots \tag{6-3}$$

In Equations 6-2 and 6-3, e is the base of the natural logarithm, n is the number of occurrences of an event - the probability of which is given by the function, λ is the positive real number, equal to the expected number of occurrences that occur during the given interval, ϕ is the real number between 0 and 1 ($\phi = 0$ is standard Poisson) called the zero inflation parameter.

To examine the distribution of this variable, SAS proc GENMOD was used to estimate two null models on all 85,277 observations and a single one for the 31,660 observations with non-zero outcomes (patients who had at least one imaging test ordered by their linked/loyal primary care doctor). Specifically the models used were simple Poisson and Zero-Inflated Poisson (ZIP) distributions for the whole data set and simple Poisson distribution on the non-zero observations. The intercept coefficients from these models and the 95% confidence intervals were used as estimators of the Poisson Lambda parameters of each proposed distribution. The dispersion for each

modeled distribution was obtained by using the PSCALE and DSCALE option in the model statement.

To visually demonstrate the relationship between the observed distribution of imaging test counts and the three proposed Poisson distributions, the parameter estimates from the null models described above were submitted to a SAS provided macro called PROBCOUNTS. This software is available for download on the public SAS support web site at the following location (checked May 2, 2010): http://support.sas.com/kb/26/161.html

The purpose of this program is succinctly described on the support web site as follows:

> The PROBCOUNTS macro computes the predicted count and the predicted probabilities of specified counts for Poisson and negative binomial models and for zero-inflated versions of these models as fit by PROC COUNTREG in SAS/ETS software and PROC GENMOD in SAS/STAT software.

A plot of the observed counts (from 0 to 15) superimposed on line graphs of the expected counts from each of the three distributions serves to compare them and will be reproduced in the Chapter 7 as Figure 7-1. The purpose of this was to visually inspect the fit of the proposed distributions with the actual imaging counts especially with respect to the 'upper tail' and the tendency of over- or under-dispersion.

To verify the estimate of Lamda(λ) for all three distributions and obtain a direct estimate of Phi(ϕ) for the ZIP distribution required a complementary approach. SAS PROC NLMIXED allows exact specification of the proposed distribution(s) and produces direct estimates for the parameters. Since an iterative approach is used, initial 'seed' values for the parameters of interest (Lam and Phi in the code fragments) are supplied. The SAS code for the two simple Poisson models is reproduced below:

```
proc nlmixed data=input.data_set;
parms Lam=1;
loglike=IMG*log(Lam)-Lam-lgamma(IMG+1);
model IMG~general(loglike);
where IMG ne 0; <<this statement is only needed for the non-zero Poisson
run;
```

The SAS code to obtain the ZIP parameters is reproduced below:

```
proc nlmixed data=input.data_set;
parms Phi=0.5 Lam=2;
if IMG=0 then prob=Phi+(1-Phi)*exp(-Lam);
if IMG=0 then loglike=log(prob);
else loglike=log(1-Phi)+IMG*log(Lam)-Lam-lgamma(IMG+1);
model IMG~general(loglike);
run;
```

Correlation between Independent Variables

Some of the clinical activity variables essentially measure the same events in slightly different ways. This is especially true for those representing outpatient visits where each of the three doctor classes (patient's own linked primary care doctor, covering primary care doctors, and specialists) is quantified by simple visit counts as well as summed RVU values. Some of the hospital activity variables also may be correlated (e.g., inpatient stays and total inpatient length of stay). Therefore, to help select among them for subsequent modeling, all 17 clinical activity variables were analyzed for (Spearman) correlation with SAS PROC CORR. The relevant results were tabulated as a correlation matrix (Table 7-1) and are discussed in the Results chapter. Spearman correlation was also performed between the two date-derived provider level variables; provider age in 2008 and provider experience in 2008.

Bivariate Relationships

Ordinary least squares regression with SAS PROC GLM was used to estimate the relationship between each independent variable and the outcome (count of imaging tests ordered by patient's loyal provider) for all observations (N=85,277). Numeric independent variables (N=28) were entered as recorded while categorical variables (N=20) were parameterized using SAS EFFECT method which yields L-1 design variables where L is number of levels in the original variable. This replicates the categorical variable coding method (reference cell) that will be used in multivariable modeling and yields degrees of freedom that are identical. Basically, this translates into an OLS linear regression with numeric variables and a one-way ANOVA with the categorical variables. For each of the 48 variables, the following parameters were obtained to measure the strength of bivariate association with the outcome: F Value, R-Squared, Correlation (R), and p value from the single variable regression output.

Variable Reduction for Modeling

Based on the correlation analysis between independent variables and evaluating the bivariate relationships with the outcome, redundant and/or collinear predictors were omitted from subsequent multivariable and multi-level modeling. The selection heuristic was to choose the one having strongest bivariate relationship with the outcome when independent variables were strongly correlated. However, for theoretical reasons, some independent variables were kept despite not having significant bivariate relationship with the outcome (e.g. provider malpractice status).

Multivariable (logistic) Modeling: Any Imaging Use

Multivariate logistic regression was used to analyze a binary outcome derived from the main outcome variable (IMG). This binary outcome variable (called ANY_IMG) is set

to ‘yes’ (N=31,660) when that patient had one or more outpatient diagnostic imaging test(s) ordered by the linked (loyal) provider and ‘no’ (N=53,617) otherwise. This was done with SAS PROC LOGISTIC with the modeled outcome set to ‘yes’. Numeric variables were entered as recorded. The categorical variables were parameterized using the SAS REFERENCE encoding method. This allows explicitly setting the reference level and results in L-1 dummy variables where L is the number of levels in the original variable. Provider level variable reference levels were as follows: Sex-Male, FMG-No, MD_Plus-No, Malpractice-No, Provider Patients-<500. Patient level reference Levels were set as follows: Sex-Male, Race-White, Insurance-Uninsured, Medications-None. The remaining patient level categorical variables were the binary clinical problem assertions and reference for each of these was set as ‘no’.

To simplify writing the model, the patient level variables may be represented by a vector P, the provider (doctor) variables by a vector D and the clinic variable(s) by C. The logistic model is expressed as Equation 6-4:

$$\text{Logit}(\rho[\text{ANY\_IMG=yes}])_i = \beta_0 + \beta_p P_{ip} + \beta_d D_{id} + \beta_c C_{ic} \qquad (6\text{-}4)$$

In Equation 6-4, i is the ith patient, p is the pth patient level variable, d is the dth doctor variable, and c is the cth clinic variable. Type III β estimates, standard errors, Chi-Squared, p-value, and odds ratios were obtained from the resulting solution output from SAS. These were used to make inferences and comparisons about joint significance and effect size of all included predictor variables. A Hosmer and Lemeshow (HL) test for goodness of fit was requested as well.

Multivariable (Poisson) Modeling: Imaging Intensity (non-zero)

This portion of the analysis seeks to determine the effect of the same predictor variables used in the logistic analysis on the intensity of imaging (given that some has occurred). Only the 31,660 observations (patients) where IMG is between 1 and 15 are used in this model. From preliminary analysis of the outcome (IMG) distribution, it was determined to be reasonably represented as a Poisson count variable. To estimate these types of models, SAS provides PROC GENMOD which allows specification of linear models with Poisson error distribution and log link function. These may be estimated using maximum likelihood methods. As above, to simplify writing the model, the patient level variables may be represented by a vector P, the provider (doctor) variables by a vector D and the clinic variable(s) by C. The Poisson model may be written as Equation 6-5.

$$\text{Log}[E(IMG_i \mid P_{ip}, D_{id}, C_{ic})] = \beta_0 + \beta_p P_{ip} + \beta_d D_{id} + \beta_c C_{ic} + e_i \qquad (6\text{-}5)$$

In Equation 6-5, i is the ith patient, p is the pth patient level variable, d is the dth doctor variable, and c is the cth clinic variable. The errors (e_i) are distributed as Poisson. After estimation, the solution output contains the coefficient estimate, standard error, Chi-Squared, and p-value for each one of the numeric variables and categorical variable levels. The model was re-estimated using DSCALE and PSCALE options to determine if the outcome distribution was over- or under-dispersed. As will be described in the Results chapter, the distribution turns out to be underdispersed, and standard error would tend to be overestimated. Therefore, a correction was NOT made for dispersion which results in somewhat conservative inferences about significance of various effect sizes.

Preparation for Multi-Level Modeling: Imaging Propensity Scores

For multi-level modeling, we wish to answer questions about variation in imaging utilization between providers (level 2) holding patient factors equal while accounting for clustering of patients (level 1) within providers. For this dissertation, the possible higher level effect of clinics as aggregations of providers will not be addressed. To simplify the specification, estimation, and interpretation of multi-level modeling results, all patient level factors were collapsed into a single risk adjusted expected imaging (propensity) variable. Previously described evaluation of the overall distribution of the outcome (IMG) determined that a zero-inflated Poisson (ZIP) distribution to be most suitable for a single model. This was done with SAS PROC GENMOD, all 85,277 observations, and the same patient level independent variables used in multivariable modeling described above. The SAS PROC GENMOD instructions were constructed so as to produce patient level predictions after the initial maximum likelihood estimation. This variable will be called IMG_PROP for 'imaging propensity' and represents the number of outpatient imaging tests that the 'average' patient with identical values of all independent variables would be expected to have. Another way of describing this technique is as regression based risk adjustment for imaging utilization. In typical provider profiling applications, these patient level predictors would be termed the 'expected' imaging utilization and summed across doctors to be compared with the 'observed' count of imaging tests actually performed on those same patients.

Multi-Level (Hierarchical) Modeling

This part of the analysis will test the relationship between a summary of each patient's imaging propensity (IMG_PROP) and imaging performed (IMG) *within* each of the 148 primary care doctor's practice. This posits that each doctor has his or her 'own'

regression equation with an intercept and slope. The interpretation of the intercept for each doctor is the ‘general’ tendency to obtain imaging on the average patient while the slope corresponds to that doctor’s ‘response’ to patients with increasing risk adjusted expectation for imaging (IMG_PROP = imaging propensity described above). With these interpretations in mind, the analysis will answer the following:

- What is the average intercept and slope of the 148 provider regression equations?
- How much do the intercepts vary from doctor to doctor?
- How much do the slopes vary from doctor to doctor?
- What is the correlation between intercepts and slopes?

Generally, in two-level hierarchical modeling, where the level 2 intercepts are allowed to vary randomly (in this study for each provider), it is important to consider whether or not to center (offset) the level 1 variable (IMG_PROP) in any way. There are three options; natural metric (no centering), grand mean (e.g., all patients), group mean (e.g., patients loyal to each doctor). Consequences of these choices affect both the value and interpretation of the intercept estimates (Raudenbush and Bryk 2002). Further, the standard errors for estimates of both fixed and random effect intercepts differ depending on choice of centering method (Luke 2004). In health services research applications, centering on the (level 2) group mean (provider in this case) is recommended (Houchens *et al.* 2007). For this study, the group means centering approach will be used. One distinct advantage is that it allows interpreting the individual provider intercepts as representing the ‘average’ tendency to order imaging while holding the expected imaging propensity (IMG_PROP) of their patients equal. On the other hand, provider intercepts estimated with the non-centered approach are clinically

less meaningful because they would have to be interpreted as the tendency to order imaging for patients with no clinical need (which numerically could be negative).

The model used will have random coefficients for intercept and slope and at the level of the patient (level 1), is written as Equation 6-6:

$$IMG_{ij} = \beta_{0j} + \beta_{1j}(IMG\_PROP_{ij} - IMG\_PROP_{.j}) + e_{ij} \quad (6\text{-}6)$$

In Equation 6-6, IMG_{ij} is the Count of outpatient images for the ith patient cared for by the jth doctor, β_{0j} is the The intercept for the jth doctor, β_{1j} is the The slope for the jth doctor, IMG_PROP_{ij} is the Imaging propensity for the ith patient cared for by the jth doctor, $IMG\_PROP_{.j}$ is the Average imaging propensity of the patients cared for by the jth doctor, and e_{ij} is the Error (disturbance) for the ith patient cared for by the jth doctor.

For simplicity let $cIMG\_PROP_{ij} = IMG\_PROP_{ij} - IMG\_PROP_{.j}$, where $cIMG\_PROP_{ij}$ is the centered imaging propensity for the ith patient cared for by the jth doctor. Therefore, our patient (level 1) model may be written as Equation 6-7.

$$IMG_{ij} = \beta_{0j} + \beta_{1j}(cIMG\_PROP_{ij}) + e_{ij} \quad (6\text{-}7)$$

The imaging utilization of each doctor's practice is characterized by two parameters: β_{0j}, the intercept for the jth doctor and β_{1j} , the slope for the jth doctor. Since the imaging propensity for each patient is centered on the mean for their doctor, the intercept is actually that doctor's mean imaging. These two parameters vary across doctors in the level-2 model as a function of the grand mean and random disturbance such that $\beta_{0j} = \gamma_{00} + u_{0j}$ and $\beta_{1j} = \gamma_{10} + u_{1j}$, where β_{0j} is the intercept (mean imaging) for the jth doctor, γ_{00} is the average of the doctor means of imaging use across the population of doctors, u_{0j} is the individual variation from the average intercept for the jth doctor, β_{1j} is the slope for the jth doctor, γ_{10} is the average imaging propensity / imaging

utilization regression slope across the doctors, and u_{1j} is the individual variation from the average slope for the jth doctor

Combining the two by substituting the level 2 random coefficient models into the level 1 patient level model gives Equation 6-8.

$$IMG_{ij} = \gamma_{00} + u_{0j} + \gamma_{10}\,(cIMG\_PROP_{ij}) + u_{1j}\,(cIMG\_PROP_{ij}) + \varepsilon_{ij} \quad (6\text{-}8)$$

The variance structure can be expressed as Equations 6-9 through 6-12.

$$u_{0j} \sim N(0, \tau_{00}) \quad (6\text{-}9)$$

$$u_{1j} \sim N(0, \tau_{11}) \quad (6\text{-}10)$$

$$Cov(u_{0j}, u_{1j}) = \tau_{01} \quad (6\text{-}11)$$

$$e_{ij} \sim N(0, \sigma^2) \quad (6\text{-}12)$$

In Equations 6-9 through 6-12, τ_{00} is the $Var(u_{0j})$, τ_{11} is the $Var(u_{1j})$, and τ_{01} is the the covariance between u_{0j} and u_{1j}.

Since the level 2 model has no predictors in either the intercept or slope equation, it is unconditional. Therefore, we can use multi-level regression estimates for variability in intercepts and slopes as shown in Equations 6-13 and 6-14.

$$Var(u_{0j}) = Var(\beta_{0j} - \gamma_{00}) = Var(\beta_{0j}) \quad (6\text{-}13)$$

$$Var(u_{1j}) = Var(\beta_{1j} - \gamma_{10}) = Var(\beta_{1j}) \quad (6\text{-}14)$$

For estimation with SAS PROC MIXED we divide the combined model into fixed parts (Equation 6-15) and random parts (Equation 6-16).

$$\gamma_{00} + \gamma_{10}\, cIMP\_PROP_{ij} \quad (6\text{-}15)$$

$$u_{0j} + u_{1j}\, cIMG\_PROP_{ij} + \varepsilon_{ij} \quad (6\text{-}16)$$

Simplified SAS Code is written:

proc mixed data=input.data_set;

```
class prov_id;
model IMG = cIMG_PROP;
random intercept cIMP_PROP / subject=prov_id;
run;
```

To quantify the fraction of variation in imaging utilization attributable to patients and providers it is necessary to obtain patient level residuals ($e_{ij} \sim \sigma^2$) from reduced models and the null model. These can be represented as follows: Patient level Imaging Propensity only (Equation 6-17), Patient level Imaging Propensity and provider slope only (Equation 6-18), Patient level Imaging Propensity and provider intercept only (Equation 6-19), and Null Model (Equation 6-20):

$$IMG_{ij} = \gamma_{00} + \gamma_{10}\,(IMG\_PROP_{ij}) + \varepsilon_{ij} \tag{6-17}$$

$$IMG_{ij} = \gamma_{00} + \gamma_{10}\,(cIMG\_PROP_{ij}) + u_{1j}\,(cIMG\_PROP_{ij}) + \varepsilon_{ij} \tag{6-18}$$

$$IMG_{ij} = \gamma_{00} + u_{0j} + \gamma_{10}\,(cIMG\_PROP_{ij}) + \varepsilon_{ij} \tag{6-19}$$

$$IMG_{ij} = \gamma_{00} + e_{ij} \tag{6-20}$$

CHAPTER 7
RESULTS

Outcome Variable Distribution

The outcome variable (prv_o_cnt=IMG) was measured for each patient (N=85,277) over two years of study and represents the count of outpatient non-invasive diagnostic imaging tests ordered by the patient's linked (loyal) primary care provider. Univariate statistics and the distribution were shown (Table 5-4 and Figure 5-1 respectively) in Chapter 5. To estimate the Poisson Lambda for the whole sample (N=85,277 including the 53,617 zero observations), SAS PROC GENMOD was used with a null model (IMG = / dist=poi). Lambda was thus estimated at 0.71459 with 95% confidence interval of (0.70718-0.72200). The dispersion (extent to which the observed variance exceeds the expected) was estimated by Deviance / Degrees of Freedom (145443.64 / 85,276) at 1.7056 which means that there are more observations at higher values of n (3-15) than expected by a Poisson distribution with Lambda=0.71459. Such 'overdispersion' can result in underestimation standard errors in subsequent modeling.

Also using SAS PROC GENMOD and a null model (IMG = / dist=poi), a second estimation, using only the non-zero observations, of Poisson Lambda and dispersion was made in the same manner as described above. In this case, the Poisson Lambda was estimated at 1.92476 with 95% confidence interval of (1.90948-1.94005) and the dispersion was 0.7796.

Beginning with Version 9.2, SAS PROG GENMOD allowed estimation of ZIP models. In addition to specifying a model for the outcome counts, a second, so called zeromodel, was specified (IMG = / dist=zip; zeromodel =;). Using the entire sample (N=85,277) with null count and zeromodel statements, the parameters of this ZIP

distribution were estimated. The Poisson Lambda estimate was 1.4917 with 95% confidence interval of (1.4760-1.5074). The dispersion was 1.2428 and the zero inflation intercept was 0.084 with 95% confidence interval of (0.0650-0.1028). To obtain a more meaningful value for the zero inflation parameter (ϕ), the output of SAS PROC NLMIXED was used to obtain a direct estimate of 0.521 (CI: 0.516-0.526). At the same time, all three estimates of the Poisson Lambda from NLMIXED were identical to those obtained from GENMOD which is reassuring.

The output from PROBCOUNTS generated using parameter estimates obtained from the null models represent the expected distributions of the simple Poisson, non-zero Poisson and zero-inflated Poisson assumptions. They are plotted in Figure 7-1. Clearly, the best fit appears to be with the Zero-inflated Poisson distribution which militates for using it to generate risk adjusted expected imaging (propensity scores) for each patient prior to multi-level modeling seeking to characterize provider variation. However, to estimate the effect of individual patient, provider, and clinic level factors on imaging utilization the two stage (logistic followed by Poisson) approach were used. One reason is that odds ratios produced by the logistic analysis of any utilization have well understood meanings that are directly interpretable. Additionally, since the Poisson distribution of non-zero use has dispersion that is less than one, there was no need to correct effect sizes or significance levels for intensity of use. If anything, the uncorrected estimates of patient, provider, and clinic factor effects were somewhat conservative.

Correlation between Independent Variables

Several of the patient level clinical activity and other imaging utilization variables are theoretically redundant in that they essentially reflect the same phenomenon. For

example, there are four variables that derive from inpatient events; number of inpatient stays (inpt_stays), number of days in the ICU (inpt_icu_days), and the two readmission measures (inpt_read_15d and inpt_read_31d). The most closely related variable pairs are the counts versus summed RVU of visits to the loyal provider, other primary care doctors, and specialists respectively. As these are highly correlated, a choice between them (counts or summed RVU) were necessary for subsequent multivariable modeling. Also, various classes of these clinical activity counts may be correlated even though they are not measuring precisely the same activity. For example, emergency room visits and imaging tests ordered from the emergency room or inpatient stays and imaging tests ordered in the hospital are likely to be correlated. Table 7-1 lists relevant correlations between clinical activity and other imaging utilization variables. Though there are 18 separate variables (17 in the rows plus the first column=all_e_cnt), only 9 columns are shown. The missing columns had no correlations > 0.5 and were omitted for brevity.

As expected, the correlations between visit count and visit RVU for specialists (0.975), linked (loyal) provider (0.994), and covering PCP (0.978) were very high. Therefore, only one set of these would be included in multivariable modeling (visits vs RVU). Emergency room visits and total hours spent in the emergency room were also highly correlated (0.996) as were the number of inpatient stays and total days spent in hospital (0.992). The two readmission measures (15 and 31 days) were also correlated (0.857). Not included in Table 7-1 is the correlation between the age and experience level of patient's linked/loyal primary care doctor. As expected, this was quite high (0.949) and justifies using only one of them for subsequent multivariable modeling.

Bivariate relationships

The results of the individual OLS regressions are listed in Table 7-2 for patient level demographic, insurance, medications, and problem variables. For patient level clinical activity and other imaging variables, results are summarized in Table 7-3. For provider and clinic variables the results are given in Table 7-4. All of the patient level variables had highly significant linear relationships with the outcome ($p < 0.0001$). The only exception was substance abuse (binary problem variable named pr_sub) with $p=0.0014$. The magnitude of these relationships varied considerably with correlation coefficients ranging from 0.01 up to 0.39, with the strongest being with the summed RVU of visits by the patient to their loyal provider (prv_visit_rvu). Other particularly strong relationships (correlation coefficients with outcome > 0.2) were exhibited by specialist outpatient visit variables (spec_vis_count, spec_visit_rvu) and the number of active prescriptions for each patient. Finally, the unique identifier for the patient's linked (loyal) provider (prov_id on Table 7-4) had a fairly high correlation (0.2268) with the number of outpatient imaging tests ordered by that same doctor. Interestingly, both measures (count and RVU) of visit intensity to the linked (loyal) provider and specialists seemed to have stronger relationships with the outcome than the actual identity of the patient's linked (loyal) provider with all bivariate correlations being > 0.23.

Variable Reduction for Modeling

As noted above (Table 7-1), the visit counts and RVU variables were highly correlated with each other. The bivariate relationships with the outcome variable from Table 7-4 were used to guide the choice between them. In each case, the RVU version had slightly higher correlation with the outcome while all of them were significant ($p<0.0001$). Accordingly, only the three visit RVU variables (prv_visit_rvu, pcp_visit_rvu,

and spec_visit_rvu) were carried forward for further analysis. Another pair of variables with high correlation between them was the provider's age and experience. From Table 7-4, we see that the provider experience variable was significantly (p=0.0002) related to the outcome while the provider's age was not (p=0.2542). This made choosing among them straightforward: select the provider experience variable for subsequent multivariable analysis. Even though the variables representing provider foreign medical graduate (FMG) and malpractice status were not significantly related to the outcome in bivariate fashion, they were carried forward due to theoretical considerations. For subsequent modeling, the variables coding actual identity of providers (prov_id) and clinics (site_id) were omitted. However, the provider identity variable was used during the final stage of analysis: multi-level hierarchical modeling. In summary, the following variables were dropped for purposes of multivariable logistic and Poisson regression:

- er_hours (total hours in the ER)
- inpt_read_15d (count of readmit within 15 days)
- pcp_visit_count (count of outpatient visits to covering PCP)
- prv_visit_count (count of outpatient visits to loyal doc)
- spec_visit_count (count of outpatient visits to specialists)
- prov_age_08 (age in years of the provider in 2008)
- prov_id (provider identifier)
- site_id (site (clinic) identifier)

Multivariable (Logistic) Modeling: Any Imaging Use

The logistic model with all 85,277 observations, 28 patient level, 6 provider level and 1 clinic level independent variables was estimated using ANY_IMG as the outcome ('yes' when the count of imaging tests ordered by patient's linked (loyal) doctor was greater than zero and 'no' otherwise). This served to jointly test the effect of each of the 35 independent variables on whether or not the patient had any imaging ordered by their linked (loyal) doctor during the two years of study. The outcome value of 'yes' was

set to be the event/success level. Subsequent interpretation of the resulting odds ratios is such that when they are greater than one, that variable/level is associated with a higher probability of imaging.

The -2 Log Likelihood was 112,501 for intercept only and 100,678 for the full model. The 'pseudo' R-Squared was 0.13 rescaled to 0.18 and the c Statistic was 0.723. A Hosmer and Lemeshow (HL) test for lack of fit was highly significant ($p<0.0001$) with Chi-Square of 588 on 8 degrees of freedom. However, it should be noted that there is evidence that for large sample sizes (exceeding 50K as in this study) a significant HL test does not entail that a particular logistic model is useless or even poorly specified (Kramer and Zimmerman 2007, Bertolini *et al.* 2000). The only hypothesis affirmed is that there is a high probability of at least some lack of fit and with the R-Squared of 0.18, this is already established.

The individual independent variable results are listed in Tables 7-5 and 7-6 and graphically depicted in Figure 7-2. In the figure, odds ratios to the right of the reference line (1.0) imply that the variable or level was associated with an increased probability that the patient would have any imaging test during the two years of study. Note that the odds ratios for numeric variables (e.g., patient age, clinical activity variables, and provider experience) represent the increase in probability of any imaging with a unit increase in the value of that variable. For example, consider patient age. For each additional decade (from 3rd through 9th), the probability of any imaging use increased by 16%. Thus, a 90 year old patient would be about 3 times more likely to have at least one imaging test compared with a 20 year old (all else equal). On the other hand, females were only 9% more likely to have imaging than males. Recall, that

mammography has been specifically omitted from this study. Otherwise, that number would likely have been about an order of magnitude higher. Black and Hispanic patients were about 20% more likely to receive imaging as compared to whites (reference level).

For the most part, insurance status was either not significant or had a small effect size in the expected direction (greater likelihood of any imaging compared with reference of uninsured). The notable exception was Medicare with a significant (<0.0001) and substantial negative effect size (OR = 0.752, ~25% less likely to have any imaging than uninsured--or self-pay--patients and even greater compared with other insurance types). The only other significant insurance type was Managed and patients were ~18% more likely to have imaging compared with uninsured. The Blue Cross group (BCBS) was marginally significant with a (~14%) positive effect on imaging use.

The number of medications patients were taking had no effect on imaging use. Individual binary clinical problems were significant in six out of nine instances. There was only one clinical problem that seemed to increase likelihood of imaging and that was Trauma (OR=1.24). The other five problems were associated with decreased imaging when present: Cancer (OR=0.949), Congestive Heart Failure (OR=0.765), Diabetes (OR=0.729), Hypertension (OR=0.784), and Substance Abuse (OR=0.793). The count variable which subsumed the remaining clinical problem list entries not categorized above (Other Problems) showed a significant (though small) positive effect on imaging (OR for 1 unit increase = 1.019). At the median level of 6.0, this would result in ~12% increase in likelihood of any imaging compared with none.

Clinical activity variables tested included the summed RVU of outpatient visits. As expected, visits to the patient's linked (loyal) doctor were strongly related to imaging

(prv_visit_rvu: OR=1.164 for a 1 unit increase). The maximum value for this variable is 17 RVU. Thus, patients having visits to their linked (loyal) doctor over two years totaling 17 RVU would be at least 13 times more likely to have imaging compared with those having a single visit with fractional RVU. The variable representing visits to other primary care doctors (pcp_visit_rvu) was not significant. However more visits to specialists (spec_visit_rvu: OR=1.025 for 1 unit increase) were associated with higher likelihood of imaging (ordered by the patient's linked/loyal doctor). Of the six clinical activity variables that measured hospitalization, only two were significantly associated with primary care imaging. These were 24 hour observation admissions (obs_stays: OR=1.07 for 1 unit increase) and the total inpatient length of stay (inpt_los_total: OR=0.987 for a 1 unit increase). The seeming discrepancy makes some sense by speculating that short stays for observation might indicate and/or engender need for imaging that would be performed later (as an outpatient).

Two of the additional imaging utilization variables had small effect sizes: outpatient ordered by specialists (spec_o_cnt: OR=1.025) and inpatient (all_i_cnt: OR=1.020). Emergency room imaging (all_e_cnt) was not significant. The number of outpatient images ordered by other primary care doctors (pcp_o_cnt: OR=1.157) was positively associated with likelihood of imaging by the patient's own (loyal) doctor. For future applications using different data sources (risk adjustment for provider profiling of imaging utilization) these additional imaging variables can probably be omitted with little consequence because the remaining clinical activity variables will capture the same phenomenon. Clearly, hospital events are, by definition, correlated with associated imaging (e.g., ER visits and imaging performed in the ER). Likewise, visits to specialists

and other (covering) primary care doctors could stand in for the outpatient imaging ordered by these same doctors (e.g., omitting pcp_o_cnt might allow pcp_visit_rvu to become significant).

Turning to the provider and clinic level variables (Table 7-6) we see that the amount of experience is negatively associated with likelihood to obtain imaging (OR=0.997 for each additional year). However the effect size is rather small. Considering that the range of experience was 5-50 years, this implies that likelihood of any imaging decreases by only about 10% between least and most experienced doctors. The gender of the doctor has a greater effect than experience, with women (OR=1.14) being 14% more likely to order imaging on their patients compared with males. Foreign medical graduates (FMG: OR=1.11) and doctors with additional academic credentials beyond M.D. (MD_Plus: OR=1.37) tend to order tests on more of their patients than American-trained and M.D. only primary care doctors. Malpractice (whether or not the doctor has been sued in past 10 years) has no significant effect. It is certainly possible that the small 'event rate' for the malpractice variable (Number Yes=7/148) contributes to the lack of significance. However, the other two (significant) provider variables also had small numbers of yes/true values (FMG=8/148, MD_Plus=14/148).

The two variables measuring practice size were both positively related to likelihood of imaging. For the (categorical) number of patients in each provider's panel, the three levels that were greater than reference (<500) had from 10-16% more likelihood to obtain imaging. The number of active providers practicing in each of the 15 clinics was

slightly positively associated with greater tendency for assigned patients to get imaging (OR=1.014) which translates into a 20% increase over the range (5-18).

Multivariable (Poisson) Modeling: Imaging Intensity (Non-zero)

The Poisson model on the 31,660 observations with non-zero outcomes had Deviance/df of 0.6613 and Chi-Squared/df of 0.7935. When the same model was run with the DSCALE and PSCALE options the resulting scale parameters were 0.8132 and 0.8908 respectively. This implies that the Poisson distribution of the outcome (number of outpatient images ordered by the linked/loyal doctor) for these non-zero patients is *underdispersed* and that the standard errors for the coefficients might tend to be slightly overestimated. Therefore, inferences about the significance of variables and levels would, if anything, be conservative and can be discussed with some confidence. These uncorrected coefficients and standard errors are presented in Tables 7-7 and 7-8, and Figure 7-3 displays these same results in terms of 95% Wald confidence intervals.

When discussing the parameter coefficient estimates in terms of effect size and direction on imaging intensity, it is important to recall that the link function for the Poisson model was log rather than linear. This means that we can't translate the value of the estimate into an additive number of imaging tests per patient for the variable or level in question. However after exponentiation (far right columns in Tables 7-7 and 7-8), the values can be interpreted as (multiplicative) incident rate ratios (IRR). To keep these in perspective, recall that the intercept for a null Poisson model on the 31,660 non-zero observations is 0.6548 which after exponentiation is 1.925. As noted above, this is the Poisson Lambda for the non-zero imaging counts and can be interpreted as the expected number of images that the 'average' (non-zero) patient would have in two years.

For the categorical variables, using reference cell encoding, the interpretation of the exponentiated coefficients is straightforward. For example, females have about 1.027 times more imaging tests than males (all else equal). Therefore, for the average female patient that had imaging, the expected count would be 1.925 times 1.027 or 1.977 images over two years. For the numerical variables the exponentiated estimate represents a multiplier applied for each additional unit value. Thus for patient age over the range of 17-103 years there is an 86 year difference which translates into Exp(86*0.0053) = 1.58 times more imaging tests between the youngest and oldest patients (all else equal). Moving on to race, we note that Black and Hispanic patients tend to have more imaging tests than whites (the reference level). With patient insurance category, only Medicare is significant compared with uninsured/self-pay with IRR of ~0.92. Recalling the logistic results, we can say that patients with Medicare are less likely to have any imaging and those that do tend to have a lower number of imaging tests than patients in other payer categories.

As with the logistic analysis for any imaging use, number of active outpatient medications had no effect on number of imaging tests. With three exceptions, the binary clinical problem variables were not significant. These were chronic renal failure (CRF), diabetes, and hypertension. All else equal, patients with these clinical problems tended to get fewer imaging tests than those without them. The variable representing the count of other clinical problems was significantly and positively associated with a larger number of imaging tests as the problem count increased (IRR=1.003 per additional problem).

The clinical activity variable most strongly associated with the number of imaging tests was the summed RVU of visits to the linked (loyal) primary care doctor. Over the range of this variable, patients with 17 RVU worth of visits to their linked primary care doctor had about 1.6 times more imaging tests than those with a single visit (totaling less than one RVU). Visits to specialists were also positively associated with number of imaging tests, though much less strongly with IRR for each additional RVU of 1.006 compared with 1.029. On the other hand, when patients saw other (covering) primary care doctors, the number of imaging tests ordered by their own (linked/loyal) doctor was slightly lower. The only other clinical activity variable that was significantly associated with the number of imaging tests was the total inpatient length of stay which had a negative effect. For example, one patient spent a total of 179 days in the hospital over the two years of study. That person would be expected to have less than 30% (Exp(179*-0.007)=0.29) of the number of imaging tests as a patient that had not been in hospital at all during the study. This may seem counter intuitive at first, but consider that patients who spend many days in hospital have more imaging tests performed there. Since the results of these tests are available to the patient's primary care doctor, he or she would likely find the answers to their diagnostic questions in tests already performed and not need to order new ones in the outpatient setting. Lastly, the variables representing outpatient imaging tests ordered by other doctors (specialists and covering primary care doctors) and while in the hospital were positively associated with the number of outpatient imaging tests ordered by the patient's linked (loyal) doctor.

Focusing on provider and site level variables (Table 7-8) and recalling the any imaging (logistic) results, a similar pattern emerges for the provider experience and

gender variables. Over the observed range of provider experience (5-50 years) the most experienced clinicians ordered about 87% as many tests as the least experienced. Female physicians ordered more tests than males but the difference was only about 7%. Doctors with additional training after their M.D. ordered about 10% more imaging. As with the any imaging (logistic) analysis, the provider's malpractice status had no effect on number of images ordered. The two middle practice size categories (500-799, 800-1000) had significant but small positive effects on number of images compared with reference (<500). The final provider variable, foreign medical graduate status (FMG_Yes), had a significant (p=0.023) negative effect (4%) on number of images per patient whereas the same variable was associated with a greater likelihood (11%) of ordering any imaging (logistic) on a given patient. Finally, the clinic size (number of doctors) variable had a small but significant positive effect on the number of imaging tests ordered on assigned patients. Over the range of practice sizes of 5-18 doctors, this translates into about 4% more imaging tests.

Comparison of Any Imaging and Imaging Intensity Results

To compare the effect of the various independent variables and levels on any imaging use (logistic) and imaging intensity (non-zero Poisson) it is useful to plot the respective odds ratios and coefficients. Figure 7-4 shows odds ratios and coefficients for the 32 variables / levels that were significant for either any imaging use or imaging intensity. Figure 7-5 shows the 9 variables that had relatively small effect sizes (odds ratio near one, coefficient near zero) with the axes scaled down to show the relationships to better advantage. These were all numeric rather than categorical variables which explains the small effect sizes (for a unit change in value) that are still quite significant. Of these, only one (summed RVU of visits to covering primary care

doctors) was significant for imaging intensity (non-zero Poisson) but not for any imaging use (Logistic).

In both Figures 7-4 and 7-5, note that the general tendency is for variables and/or levels to be concordant with respect to their effect size and direction (when both are significant) between any imaging use (logistic odds ratio, X Axis) and imaging intensity (Poisson regression coefficient, Y Axis). There is only a single exception and this is with the variable coding for whether or not the linked (loyal) primary care doctor is a foreign medical graduate (FMG). It seems that FMG primary care doctors were more likely to order some imaging but ordered fewer imaging tests when they did so.

Preparation for Multilevel Modeling: Imaging Propensity Scores

Patient level predictions from the zero inflated Poisson (ZIP) model using all 28 patient level variables for both the count and zero-model portions were calculated, called IMG_PROP, and stored in a new data set along with the original outcome (IMG) and the (coded/anonymous) provider ID of that patient's linked primary care doctor. Table 7-9 compares the raw outcome (IMG) and the predictions from the ZIP model (IMG_PPOP).

The mean of IMG_PROP for each provider was subtracted from the original value for each patient to form a centered imaging propensity variable (cIMG_PROP), which can be expressed as Equation 7-1.

$$cIMG\_PROP_{ij} = IMG\_PROP_{ij} - IMG\_PROP_{.j} \qquad (7\text{-}1)$$

In Equation 7-1, i is the ith patient, j is the jth provider, and .j is the average for the jth provider.

The result of the above described centering operation is illustrated by plotting the raw imaging propensity scores (IMG_PROP) and centered imaging propensity scores

(cIMG_PROP) against the mean outcome (images per patient) for each provider. These are shown in Figures 7-6 and 7-7 respectively.

Multi-Level (Hierarchical) Modeling

The SAS PROC MIXED procedure on the full two level model completed in ~20 seconds and converged after 4 iterations. Results from SAS are reproduced in Tables 7-10 through 7-15.

These results are summarized in Table 7-16 in terms of fixed (patient level) effects and random (patient and provider levels) variance.

The fixed effect intercept (0.7171) is the average doctor's mean imaging. This is nearly identical to the raw mean value of 0.7146 obtained by dividing the total number of imaging tests ordered by the patient's linked provider (N=60,938) by the number of patients in the whole study cohort (N=85,277). A (95%) range of plausible values for doctor's mean imaging (intercept) around 0.7171 can be constructed using the variance (0.0835) by $0.7171 \pm 1.96(0.0835)^{1/2}$ which gives (0.151, 1.283).

Note that this is rather wider than the 95% confidence intervals (0.6695, 0.7648) on the estimate of the fixed intercept provided by SAS, which used the standard error.

The fixed effect of cIMG_PROP (0.9919) is interpreted as the average doctor's response (slope) in number of imaging tests ordered for a unit change in the imaging propensity score (cIMG_PROP). The fact that this is very close to 1.0 implies that the scale of the imaging propensity score is correct (at least around the mean value). In other words, on average, as the expected amount of imaging increases by one 'unit', actual imaging utilization increased by one extra test per patient. As with the intercept,

we can calculate a (95%) range of plausible values for the slopes using the variance (0.1567) by $0.9919 \pm 1.96(0.1567)^{1/2}$ which gives (0.216, 1.768).

As with the intercept, the plausible range of the slope is wider than the SAS calculated 95% confidence interval on the slope parameter estimate (0.925, 1.058).

Intraclass correlations (ICC) for intercept and slope were obtained by dividing each component variance (τ_{00} for intercept and τ_{11} for slope) by the residual variance (σ^2) plus itself. For intercept, this is 0.0835 divided by (0.0835+1.1934) which gives 0.065. For slope, the ICC is 0.1567 divided by (0.1567+1.1934) which gives 0.116. This implies that about 6% of the variance in intercepts is between doctors and about 12% of the variance in slopes is between doctors. It is helpful to recall that an ICC of 0 would mean that all doctors exhibit the same IMG/cIMG_PROP relationship and clustering of patients by doctor had no effect (i.e., the hierarchical modeling not informative). On the other hand, an ICC approaching 1 would mean that any given doctor's patients have nearly identical adjusted imaging utilization and very small variation between them.

Model based estimates (including standard errors and 95% confidence intervals) of individual provider intercept and slope were obtained by requesting the solution for the random portion of the model. Reliability for the individual provider intercept estimates can be calculated as the overall provider intercept variance (0.1567) divided by the provider's own variance (standard error squared) and the overall reliability for provider intercept estimates is the average of our 148 doctors which is 0.965. Similarly, the individual reliability for each provider's slope estimate is the overall variance (0.0835) divided by the individual variance (standard error squared) with the aggregate reliability being the average of these for the 148 doctors which is 0.939. The high

reliability of both the intercept and slope estimates is reassuring and supports interpreting them as representing each provider's mean tendency to order imaging tests (intercept) and their response (slope) to patient level imaging propensity represented by the cIMG_PROP (risk adjusted expected imaging) variable.

The correlation between each doctor's general tendency to order imaging (intercept) and his or her response to patient imaging propensity (slope) can be expressed as $\rho(\beta_{0j}, \beta_{1j})$ which is estimated by $\tau_{01}/(\tau_{00}\ \tau_{11})^{1/2}$. Substituting from Table 7-16 gives $0.0810 / (0.0835 \times 0.1567)^{1/2}$ which turns out to be 0.7081. This implies a substantial correlation between the average tendency to use imagining, and the increase in the number of images providers order on their patients with higher imaging propensity (i.e., sicker). A scatter plot of the intercepts (X axis) and slopes (Y axis) for all 148 providers is shown in Figure 7-8 and serves to visualize the relationship between them. The quadrants in this slope versus intercept plot are labeled A-D and are further detailed in Tables 7-17 and 7-18.

Less than 20% of the providers have discordance between their slopes and intercepts (quadrants A and D) with the remaining 80% (quadrants B and C) being concordant with respect to their tendency to image and their response to patient imaging 'need' as represented by the imaging propensity variable. Therefore, in general, if a primary care doctor tends to order imaging *less than average*, odds are 4.5:1 that he or she will also increase the amount of images that they obtain on 'sicker' patients *less than average*. For example, provider C.C. obtained about 29 images per 100 patients over the course of the study and increased their image utilization by about 15 images per 100 for each unit increase in imaging propensity. It is useful to plot imaging

propensity versus actual imaging utilization for all the patients cared for by Doctor C.C. and this is shown in Figure 7-9.

Likewise, if a doctor tends to order imaging *more than average*, odds are 3.7:1 that he or she will increase the amount of images that they order on 'sicker' patients *more than average*. For example, provider B.B. obtained about 199 images per 100 patients over the course of the study and increased their image utilization by about 264 images per 100 for each unit increase in imaging propensity. The observed imaging versus imaging propensity for Dr. B.B. is plotted in Figure 7-10.

The discordant doctors (quadrants A and D in Figure 7-8) are not only few in number but tend to cluster near the mean value of the slope (with standard errors overlapping the average slope of zero) such that only three doctors are significantly discordant with one having low intercept / high slope and two having high intercept / low slope. For example, the provider labeled A.A. in the scatterplot obtained about 50 images per 100 patients over the two years of study and increased imaging by about 119 images per 100 for every unit increase in imaging propensity. Another way of saying this is that Doctor A.A. has a somewhat higher threshold for obtaining imaging on any given patient but tends to order more images as patient need for imaging increases.

In contrast, provider D.D. obtained about 85 images per 100 patients and increased their imaging by about 58 images per 100. One might speculate that provider D.D. is generous with imaging in general but less discriminating in terms of increasing utilization according to patient need. On the other hand, the two quadrant D providers (D.D. and the one just above on the scatter plot) may be relatively liberal in terms of

both imaging and referral to specialists. This might mean that their sicker patients get less images ordered by the primary care provider because they tend to refer at a lower threshold and at least some images would be ordered by the specialists rather than themselves.

Another interesting visualization is to plot the only the intercepts (provider's mean imaging) and 95% confidence intervals sorted according to practice site, and this shown in Figure 7-11. Similarly, a plot of the slopes (provider's change in imaging as imaging propensity increases) with 95% confidence intervals is shown as Figure 7-12. One important observation from the provider slopes plot (Figure 7-12) is that *all providers* increase their diagnostic image ordering in response to additional patient need (none of the scaled imaging vs. propensity slopes are below zero). The alternative is that some doctor's slopes could be negative such that the amount of imaging actually decreased for sicker patients. One explanation for this (counterfactual) would be that those with negative slopes, refer sicker patients to specialists who themselves obtain the needed imaging tests. The fact that this did NOT occur at all in the current study implies that even sick patients who see many specialists continue to have at least some of their care rendered by the linked (loyal) primary care provider. This should come as no surprise given that demonstration of this 'loyalty' relationship was the main inclusion criteria for both patients and providers.

The reduced level 1 model (IMG = IMG_PROP) yielded a patient level residuals (error variance) of 1.3061 while the null model (IMG=;) gave patient level residuals (error variance) of 1.5782. These are combined with the error variance from the full two level model (1.1934) in Table 7-19.

The overall amount of explained variation in outpatient imaging utilization after accounting for provider ID and all the patient level variables (risk-adjustment) is 24.4% (0.3848/1.5782). Of that, roughly 70% (0.2721/0.3848) is attributable to patient level factors as captured in the imaging propensity variable and the remaining 30% is attributable to provider variation. One implication is that about three quarters of the variation in the number of outpatient imaging tests ordered by a primary care doctor on loyal patients is unexplained. This is despite taking into account a robust and large set of patient factors as well as all between doctor differences in imaging utilization habits (by directly modeling unique provider identity). Of the roughly 25% of variation in primary care outpatient imaging utilization that can be explained, the majority (70%) is attributed to factors that mostly relate to each patient's clinical 'need' for imaging, regardless of who their primary care doctor is. The remaining 30% arises from differences in the tendency for primary care doctors to order imaging which may be partitioned into intercept (~10%) and slope (~20%) components. The next chapter will summarize and discuss these results in terms of advances in knowledge, applications in imaging utilization management and provider profiling. Directions for future research with these (and similar) data sources as well as some policy implications will be covered as well.

Table 7-1. Spearman correlations between clinical activity and other imaging variables.

Variable Name	all_e_cnt	all_i_cnt	spec_o_cnt	spec_visit_count	pcp_visit_count	prv_visit_count	er_visits	inpt_stays	inpt_read_15d
all_i_cnt	0.450								
spec_o_cnt	0.226	0.273							
pcp_o_cnt	0.095	0.056	0.107						
spec_visit_count	0.260	0.278	0.574*						
pcp_visit_count	0.111	0.048	0.132	0.195					
prv_visit_count	0.253	0.230	0.241	0.342	0.125				
spec_visit_rvu	0.260	0.279	0.575*	0.975*	0.194	0.348			
pcp_visit_rvu	0.115	0.053	0.135	0.199	0.994*	0.128			
prv_visit_rvu	0.254	0.239	0.249	0.351	0.112	0.978*			
er_visits	0.843*	0.423	0.228	0.272	0.131	0.263			
er_hours	0.857*	0.429	0.229	0.273	0.130	0.267	0.996*		
obs_stays	0.295	0.148	0.250	0.298	0.076	0.136	0.290		
inpt_stays	0.469	0.726*	0.303	0.331	0.066	0.231	0.455		
inpt_los_total	0.474	0.742*	0.299	0.327	0.066	0.232	0.459	0.992*	
inpt_icu_days	0.203	0.366	0.101	0.110	0.011	0.091	0.184	0.283	
inpt_read_15d	0.243	0.343	0.106	0.121	0.028	0.099	0.231	0.302	
inpt_read_31d	0.283	0.398	0.126	0.141	0.030	0.116	0.267	0.352	0.857*

Key To Variable Names Is Below

all_e_cnt	count of images done in ER
all_i_cnt	count of images done as inpatient
spec_o_cnt	count of outpatient images ordered by specialists
pcp_o_cnt	count of outpatient images ordered by covering PCP
spec_visit_count	count of outpatient visits to specialists
pcp_visit_count	count of outpatient visits to covering PCP
prv_visit_count	count of outpatient visits to loyal doc
spec_visit_rvu	sum of RVU of outpatient visits to specialists
pcp_visit_rvu	sum of RVU of outpatient visits to covering PCP
prv_visit_rvu	sum of RVU of outpatient visits to loyal doc
er_visits	count of ER visits
er_hours	total hours in the ER
obs_stays	count of observation stays
inpt_stays	count of inpatient stays
inpt_los_total	total days in hospital
inpt_icu_days	total days in ICU
inpt_read_15d	count of readmit within 15 days
inpt_read_31d	count of readmit within 31 days

NOTE: For brevity, columns where ALL correlations were < 0.5 are omitted.
* Correlations above 0.5.

Table 7-2. Bivariate relationship between patient level variables and outcome (imaging counts).

Type	Description	Variable Name	F Value	R-Squared	Correlation	p value
Categorical	patient identified race	Race	29.98	0.0011	0.0324	<0.0001
Categorical	patient sex	Sex	258.44	0.0030	0.0550	<0.0001
Numeric	patient age in 2008	age_08	3087.75	0.0349	0.1869	<0.0001
Categorical	patient's payer of record in 2008	PayerGroup	343.33	0.0236	0.1536	<0.0001
Categorical	Active prescriptions in 2008	meds_cat	1313.04	0.0442	0.2101	<0.0001
Binary	coronary artery disease	pr_cad	445.63	0.0052	0.0721	<0.0001
Binary	cancer	pr_can	476.77	0.0056	0.0746	<0.0001
Binary	congestive heart failure	pr_chf	123.35	0.0014	0.0379	<0.0001
Binary	chronic renal failure	pr_crf	132.79	0.0016	0.0394	<0.0001
Binary	diabetes	pr_dm	436.64	0.0051	0.0713	<0.0001
Binary	obesity	pr_obs	150.85	0.0018	0.0421	<0.0001
Binary	hyptertension	pr_htn	628.14	0.0073	0.0855	<0.0001
Binary	substance abuse	pr_sub	10.19	0.0001	0.0110	0.0014
Binary	trauma	pr_trm	48.85	0.0006	0.0239	<0.0001
Numeric	count of active problems other than those listed	oth_prb	3115.43	0.0353	0.1877	<0.0001

Table 7-3. Bivariate relationship between clinical activity variables and outcome (imaging counts).

Type	Description	Variable Name	F Value	R-Squared	Correlation	p value
Numeric	total hours in the ER	er_hours*	1323.13	0.0153	0.1236	<0.0001
Numeric	count of ER visits	er_visits	1412.2	0.0163	0.1276	<0.0001
Numeric	total days in ICU	inpt_icu_days	78.36	0.0009	0.0303	<0.0001
Numeric	total days in hospital	inpt_los_total	753.11	0.0088	0.0935	<0.0001
Numeric	count of readmit within 15 days	inpt_read_15d*	214.52	0.0025	0.0501	<0.0001
Numeric	count of readmit within 31 days	inpt_read_31d	273.09	0.0032	0.0565	<0.0001
Numeric	count of inpatient stays	inpt_stays	1562.36	0.0180	0.1341	<0.0001
Numeric	count of observation stays	obs_stays	647.26	0.0075	0.0868	<0.0001
Numeric	count of images done in ER	all_e_cnt	1281.08	0.0148	0.1217	<0.0001
Numeric	count of images done as inpatient	all_i_cnt	663.8	0.0077	0.0879	<0.0001
Numeric	count of outpatient images ordered by covering PCP	pcp_o_cnt	581.73	0.0068	0.0823	<0.0001
Numeric	count of outpatient images ordered by specialists	spec_o_cnt	2729.8	0.0310	0.1761	<0.0001
Numeric	count of outpatient visits to covering PCP	pcp_visit_count*	436.63	0.0051	0.0713	<0.0001
Numeric	sum of RVU of outpatient visits to covering PCP	pcp_visit_rvu	536.26	0.0063	0.0791	<0.0001
Numeric	count of outpatient visits to loyal doc	prv_visit_count*	14540.96	0.1457	0.3817	<0.0001
Numeric	sum of RVU of outpatient visits to loyal doc	prv_visit_rvu	15228.91	0.1515	0.3893	<0.0001
Numeric	count of outpatient visits to specialists	spec_visit_count*	4786.98	0.0532	0.2305	<0.0001
Numeric	sum of RVU of outpatient visits to specialists	spec_visit_rvu	5683.23	0.0625	0.2500	<0.0001

* Variables NOT carried forward to multivariable analysis.

Table 7-4. Bivariate relationship between provider and clinic level variables and outcome (imaging counts).

Type	Description	Variable Name	F Value	R-Squared	Correlation	p value
Categorical	whether provider has been sued in last 10 years	mp_flag	1.05	0.0000	0.0032	0.3054
Categorical	whether provider is foreign medical graduate	prov_fmg	0.17	0.0000	0.0000	0.6761
Categorical	whether provider has a degree beyond MD	prov_md_plus	32.71	0.0004	0.0195	<0.0001
Categorical	number of patient's in provider practice in 2008	prov_pat_cat	79.4	0.0028	0.0528	<0.0001
Categorical	provider sex	prov_sex	66.9	0.0008	0.0279	<0.0001
Numeric	age in years of the provider in 2008	prov_age_08*	1.3	0.0000	0.0045	0.2542
Numeric	number of years after provider MD graduation in 2008	prov_exp_08	14.37	0.0002	0.0130	0.0002
Identifier	anonymous provider identifier	prov_id*	31.43	0.0514	0.2268	<0.0001
Numeric	number of doctors actively practicing at the clinic in 2008	site_docs	189.38	0.0022	0.0471	<0.0001
Identifier	site (clinic) identifier	site_id*	100.16	0.0162	0.1272	<0.0001

* Variables NOT carried forward to multivariable analysis.

Table 7-5. Patient level results from multivariable logistic model on any imaging use.

Type	Variable / Level	Estimate	Standard Error	Chi-Square	Pr > ChiSq	Odds Ratio
Demographics	Age	0.016	0.001	483.106	<0.0001	1.016
	Sex F	0.083	0.018	22.011	<0.0001	1.087
Race	Black	0.212	0.035	36.472	<0.0001	1.236
	Hispanic	0.232	0.032	52.404	<0.0001	1.261
	Other	0.091	0.029	9.695	0.0018	1.095
Insurance	BCBS	0.133	0.067	3.933	0.0473	1.143
	Commercial	0.092	0.071	1.712	0.1908	1.097
	Managed	0.168	0.068	6.065	0.0138	1.183
	Medicare	-0.286	0.070	16.570	<0.0001	0.752
	State	0.028	0.072	0.150	0.6982	1.028
	Other	0.016	0.078	0.041	0.8390	1.016
Medications	1-5	0.044	0.037	1.381	0.2400	1.045
	6-10	0.044	0.040	1.234	0.2666	1.045
	>10	-0.017	0.046	0.134	0.7146	0.983
Problems	CAD	-0.026	0.039	0.463	0.4961	0.974
	Cancer	-0.052	0.025	4.491	0.0341	0.949
	CHF	-0.268	0.079	11.453	0.0007	0.765
	CRF	-0.113	0.069	2.650	0.1036	0.893
	Diabetes	-0.316	0.027	141.985	<0.0001	0.729
	Obesity	0.013	0.026	0.267	0.6056	1.013
	Hypertension	-0.243	0.019	160.556	<0.0001	0.784
	Substance Abuse	-0.232	0.085	7.440	0.0064	0.793
	Trauma	0.215	0.053	16.426	<0.0001	1.240
	Other (count)	0.018	0.001	193.493	<0.0001	1.019
Visits	prv_visit_rvu	0.152	0.002	4253.497	<0.0001	1.164
	pcp_visit_rvu	-0.006	0.006	1.115	0.2910	0.994
	spec_visit_rvu	0.025	0.002	276.583	<0.0001	1.025
Other Imaging	pcp_o_cnt	0.146	0.019	62.486	<0.0001	1.157
	spec_o_cnt	0.024	0.004	32.949	<0.0001	1.025
	all_e_cnt	0.005	0.009	0.282	0.5952	1.005
	all_i_cnt	0.020	0.007	8.978	0.0027	1.020
Hospital	er_visits	-0.012	0.014	0.730	0.3929	0.989
	obs_stays	0.067	0.016	17.880	<0.0001	1.070
	inpt_stays	-0.022	0.024	0.868	0.3515	0.978
	inpt_read_31d	-0.077	0.053	2.070	0.1503	0.926
	inpt_los_total	-0.013	0.004	10.709	0.0011	0.987
	inpt_icu_days	-0.013	0.015	0.736	0.3910	0.987

Reference Levels: Sex-Male, Race-White, Insurance-Uninsured, Medications-None

Table 7-6. Provider and clinic level results from multivariable logistic model on any imaging use.

Type	Variable / Level	Estimate	Standard Error	Chi-Square	Pr > ChiSq	Odds Ratio
Provider	Experience	-0.003	0.001	10.993	0.0009	0.997
	sex F	0.133	0.018	55.406	<0.0001	1.142
	FMG Yes	0.108	0.034	10.412	0.0013	1.114
	MD Plus Yes	0.318	0.029	117.859	<0.0001	1.374
	Malpractice Yes	0.020	0.040	0.262	0.6085	1.021
Provider Patients	500-759	0.123	0.028	18.621	<0.0001	1.130
	750-999	0.150	0.027	31.511	<0.0001	1.162
	1K+	0.098	0.025	14.974	0.0001	1.103
Clinic size	Active Providers	0.014	0.002	69.051	<0.0001	1.014

Reference levels: Sex-Male, FMG-No, MD_Plus-No, Malpractice-No, Provider Patients-<500

Table 7-7. Patient level results from multivariable Poisson model on imaging intensity.

Type	Variable / Level	Estimate	Standard Error	Chi-Square	Pr > ChiSq	(RR) Exp Estimate
Demographics	Age	0.0053	0.0004	183.800	<0.0001	1.005
	Sex F	0.027	0.010	8.220	0.0042	1.027
Race	Black	0.034	0.018	3.480	0.062	1.034
	Hispanic	0.073	0.016	19.500	<0.0001	1.075
	Other	-0.013	0.017	0.650	0.4198	0.987
Insurance	BCBS	-0.047	0.039	1.440	0.2305	0.954
	Commercial	-0.052	0.041	1.590	0.2071	0.949
	Managed	-0.031	0.040	0.620	0.4305	0.969
	Medicare	-0.085	0.040	4.520	0.0335	0.918
	State	-0.048	0.041	1.360	0.2435	0.953
	Other	0.000	0.045	0.000	0.9932	1.000
Medications	1-5	0.001	0.024	0.000	0.959	1.001
	6-10	0.010	0.025	0.170	0.6787	1.010
	>10	0.028	0.027	1.060	0.3028	1.028
Problems	CAD	-0.028	0.017	2.730	0.0982	0.972
	Cancer	0.013	0.012	1.200	0.2725	1.013
	CHF	-0.047	0.033	2.040	0.153	0.954
	CRF	-0.078	0.030	6.860	0.0088	0.925
	Diabetes	-0.058	0.013	21.250	<0.0001	0.944
	Obesity	0.018	0.013	2.060	0.1508	1.019
	Hypertension	-0.058	0.010	36.160	<0.0001	0.944
	Substance Abuse	0.063	0.040	2.520	0.1124	1.065
	Trauma	0.013	0.026	0.270	0.6011	1.013
	Other (count)	0.003	0.001	32.560	<0.0001	1.003
Visits	prv_visit_rvu	0.029	0.001	1479.330	<0.0001	1.029
	pcp_visit_rvu	-0.009	0.003	11.390	0.0007	0.991
	spec_visit_rvu	0.006	0.001	90.200	<0.0001	1.006
Other Imaging	pcp_o_cnt	0.026	0.008	10.740	0.001	1.026
	spec_o_cnt	0.007	0.002	18.360	<0.0001	1.007
	all_e_cnt	-0.004	0.003	1.640	0.200	0.996
	all_i_cnt	0.012	0.002	26.340	<0.0001	1.012
Hospital	er_visits	-0.005	0.005	0.790	0.3754	0.996
	obs_stays	0.006	0.007	0.870	0.3503	1.006
	inpt_stays	0.010	0.010	1.030	0.3101	1.010
	inpt_read_31d	0.023	0.021	1.200	0.274	1.023
	inpt_los_total	-0.007	0.002	17.580	<0.0001	0.993
	inpt_icu_days	-0.010	0.005	3.660	0.0558	0.990

Reference Levels: Sex-Male, Race-White, Insurance-Uninsured, Medications-None

Table 7-8. Provider and clinic level results from multivariable Poisson model on imaging intensity.

Type	Variable / Level	Estimate	Standard Error	Chi-Square	Pr > ChiSq	(RR) Exp Estimate
Provider	Experience	-0.003	0.001	28.060	<0.0001	0.997
	sex F	0.063	0.009	45.860	<0.0001	1.065
	FMG Yes	-0.041	0.018	5.150	0.0232	0.960
	MD Plus Yes	0.097	0.015	40.980	<0.0001	1.102
	Malpractice Yes	-0.008	0.021	0.140	0.7091	0.992
Provider Patients	500-759	0.044	0.015	8.540	0.0035	1.044
	750-999	0.084	0.014	35.880	<0.0001	1.087
	1K+	0.026	0.014	3.630	0.0567	1.026
Clinic size	Active Providers	0.003	0.001	12.460	0.0004	1.003

Reference levels: Sex-Male, FMG-No, MD_Plus-No, Malpractice-No, Provider Patients-<500

Table 7-9. Univariate statistics for raw imaging counts (IMG) and predictions from ZIP model (IMG_PROP).

Statistic	IMG	IMG_PROP
N	85277	85277
Minimum	0 (N=53,617)	0.0446
Maximum	15	*14.8, 15,2, 20.1
Mean	0.7146	0.7164
Standard Deviation	1.256	0.5651
Skewness	2.655	3.844
Coefficient of Variation	176	78
Sum of Observations	60938	61088
Variance	1.578	0.319
Kurtosis	10	46
Standard Error of the Mean	0.004	0.002

NOTE: The highest 3 observations (*) are shown for IMG_PROP Maximum.

Table 7-10. Dimensions.

Covariance Parameters	4
Columns in X	2
Columns in Z Per Subject	2
Subjects	148
Max Obs Per Subject	2101

Table 7-11. Estimated G correlation matrix.

Row	Effect	prov_id	Col1	Col2
1	Intercept	10562	1.0000	0.7081
2	cIMG_PROP	10562	0.7081	1.0000

Table 7-12. Covariance parameter estimates.

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	prov_id	0.08345	0.01003	8.32	<0.0001
UN(2,1)	prov_id	0.08098	0.01195	6.78	<0.0001
UN(2,2)	prov_id	0.1567	0.01953	8.02	<0.0001
Residual		1.1934	0.005789	206.13	<0.0001

NOTE: UN(1,1) = τ_{00}, UN(2,2) = τ_{11}, UN(2,1) = τ_{01}

Table 7-13. Fit statistics.

-2 Res Log Likelihood	257971.1
AIC (smaller is better)	257979.1
AICC (smaller is better)	257979.1
BIC (smaller is better)	257991.1

Table 7-14. Solution for fixed effects.

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	0.7171	0.02412	147	29.74	<0.0001
cIMG_PROP	0.9919	0.03364	147	29.48	<0.0001

Table 7-15. Type 3 tests of fixed effects.

Effect	Num DF	Den DF	F Value	Pr > F
cIMG_PROP	1	147	869.19	<0.0001

Table 7-16. Results from multi-level random coefficients model.

Fixed Effect	Symbol	Coefficient	Standard Error	t-value	p-value
Intercept	γ_{00}	0.7171	0.02412	29.74	<0.0001
cIMG_PROP	γ_{10}	0.9919	0.03364	29.48	<0.0001

Random Effect	Symbol	Variance Component	Standard Error	z-value	p-value
Patient Residual	e_{ij} (σ^2)	1.1934	0.00579	206.13	<0.0001
Provider Intercept	u_{0j} (τ_{00})	0.0835	0.01003	8.32	<0.0001
Provider Slope	u_{1j} (τ_{11})	0.1567	0.01953	8.02	<0.0001
Covariance: Intercept, Slope	Cov(u_{0j},u_{1j}) (τ_{01})	0.0810	0.01195	6.78	<0.0001

Table 7-17. Quadrants in intercept versus slope relationship plot.

Quadrant	Tendency To Order Imaging: Intercept	Response To Imaging Propensity: Slope	Number Of Providers	Percent Of Providers
A	Low (<0)	High (>0)	14	9.46
B	High (>0)	High (>0)	52	35.14
C	Low (<0)	Low (<0)	67	45.27
D	High (>0)	Low (<0)	15	10.14

Table 7-18. Exemplary providers in each quadrant.

Provider	Tendency To Order Imaging: Intercept	Response To Imaging Propensity: Slope
A.A.	-0.2176	0.1988
	(0.4995)	(1.1907)
B.B.	1.2708	1.6451
	(1.9879)	(2.637)
C.C.	-0.4268	-0.8387
	(0.2903)	(0.1532)
D.D.	0.1310	-0.4087
	(0.8481)	(0.5832)

NOTE: Numbers in parentheses for intercept are adjusted to the fixed effect mean by adding 0.7171 and numbers in parenthesis for slope are adjusted to the fixed effect mean by adding 0.9919.

Table 7-19. Comparison of null, and reduced model residuals with full model.

Patient IMG_PROP	Provider Intercept	Provider Slope	Model Name	Residual (error)	Absolute Reduction	Fraction Reduction
			Null	1.5782	-	-
	X		Provider Intercept Only	1.4996	0.0786	0.0498
X			Imaging Propensity Only*	1.3061	0.2721	0.1724
X		X	Two Level: Provider Slope	1.2727	0.3055	0.1936
X	X		Two Level: Provider Intercept	1.2389	0.3393	0.2150
X	X	X	Two Level: Provider Intercept and Slope	1.1934	0.3848	0.2438

NOTE: Fraction reduction in the residual is equivalent to R-Squared for that model.
*NON Centered. If centered, error=1.3180.

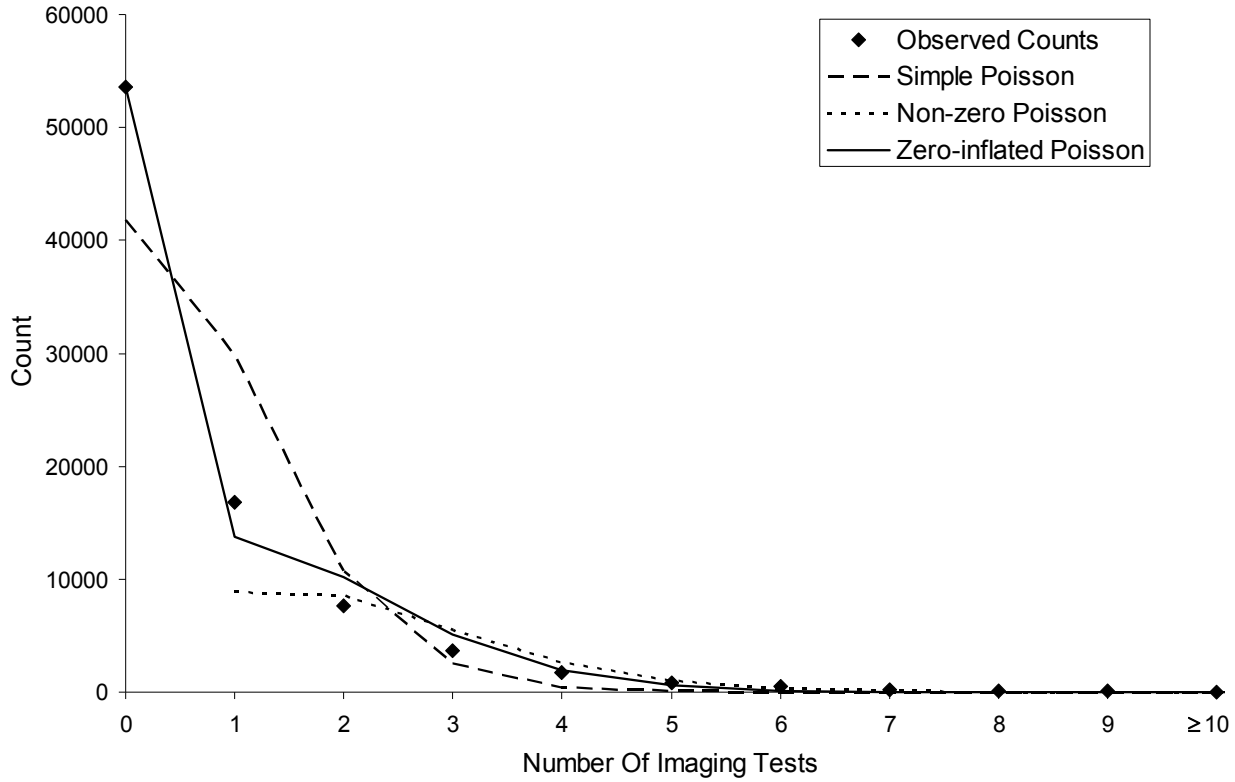

Figure 7-1. Comparison of imaging counts with three Poisson distributions.

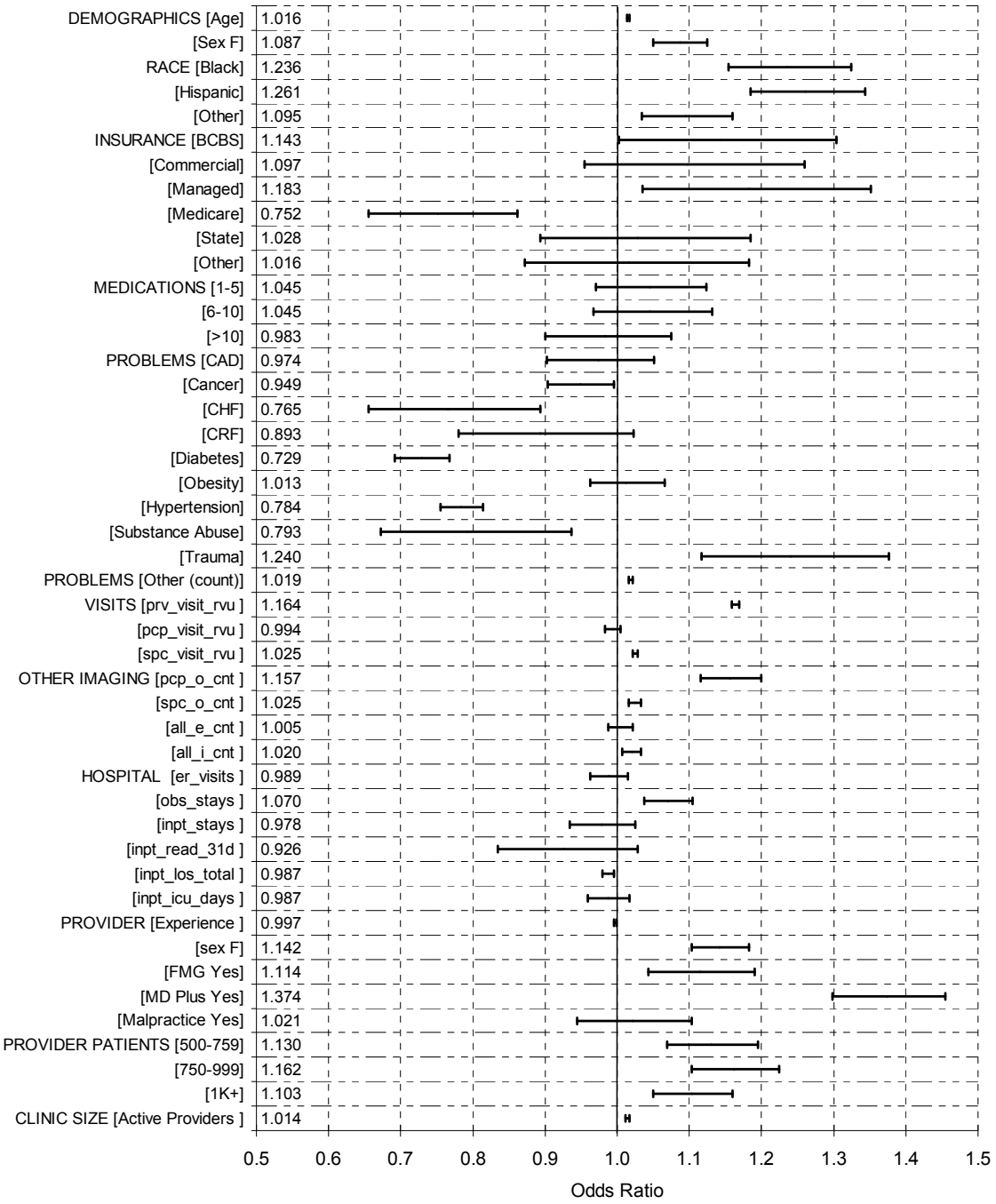

Figure 7-2. Logistic regression results for any imaging utilization. Horizontal bars represent 95% confidence intervals on odds ratio for each variable/level. Patient variable reference levels: Sex-Male, Race-White, Insurance-Uninsured, Medications-None, Problems-No. Provider variable reference levels: Sex-Male, FMG-No, MD_Plus-No, Malpractice-No, Provider Patients-<500.

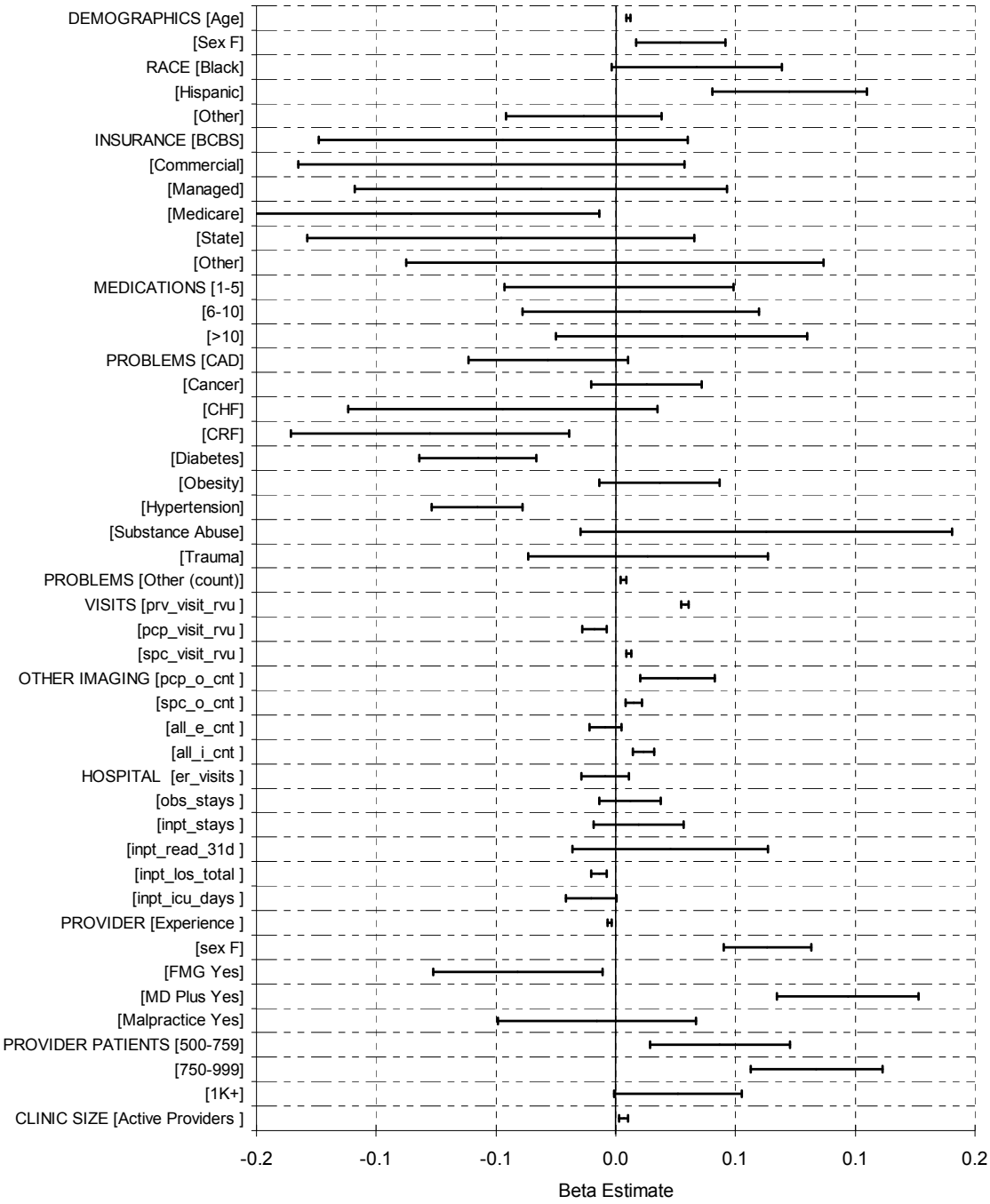

Figure 7-3. Poisson regression results for (non-zero) imaging intensity. Horizontal bars represent 95% confidence intervals on estimated coefficient for each variable/level. Patient variable reference levels: Sex-Male, Race-White, Insurance-Uninsured, Medications-None, Problems-No. Provider variable reference levels: Sex-Male, FMG-No, MD_Plus-No, Malpractice-No, Provider Patients-<500.

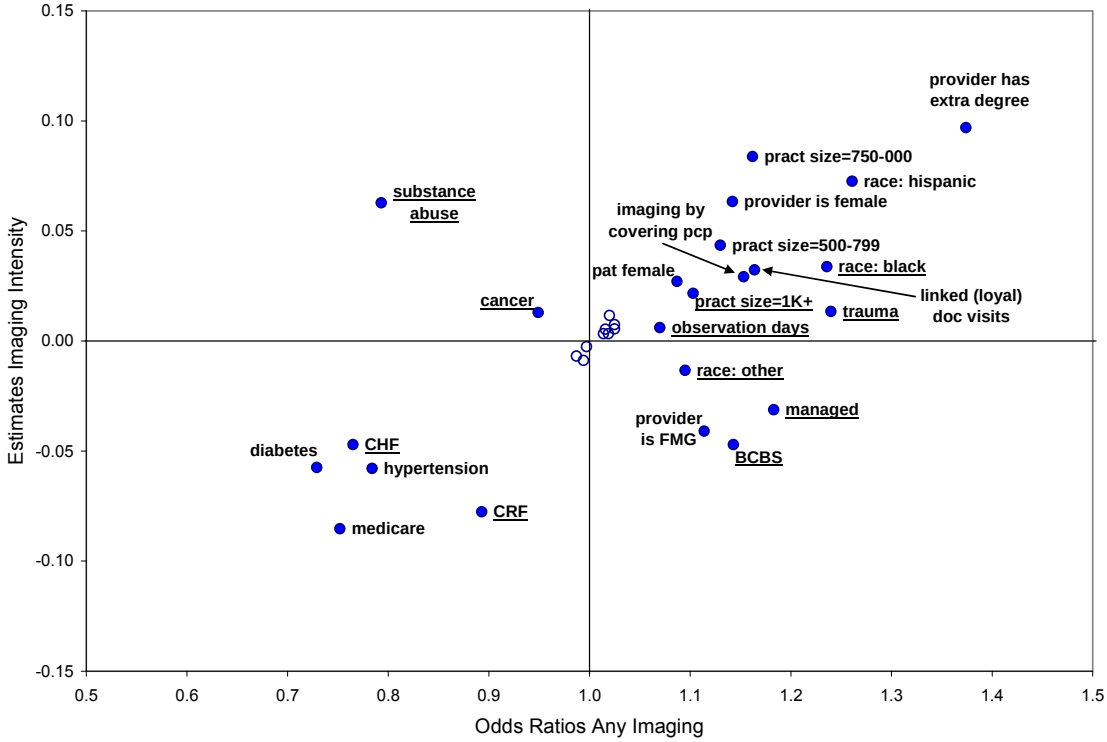

Figure 7-4. Comparison of significant variables for any imaging use and imaging intensity. Unless underlined, all variable/levels were significant for both any imaging use and imaging intensity. The underlined variable/levels were not significant for imaging intensity, except for chronic renal failure (CRF), which not significant for any imaging use. The variables indicated by open circles (near the origin) are shown again in Figure 13 with appropriate axis scaling.

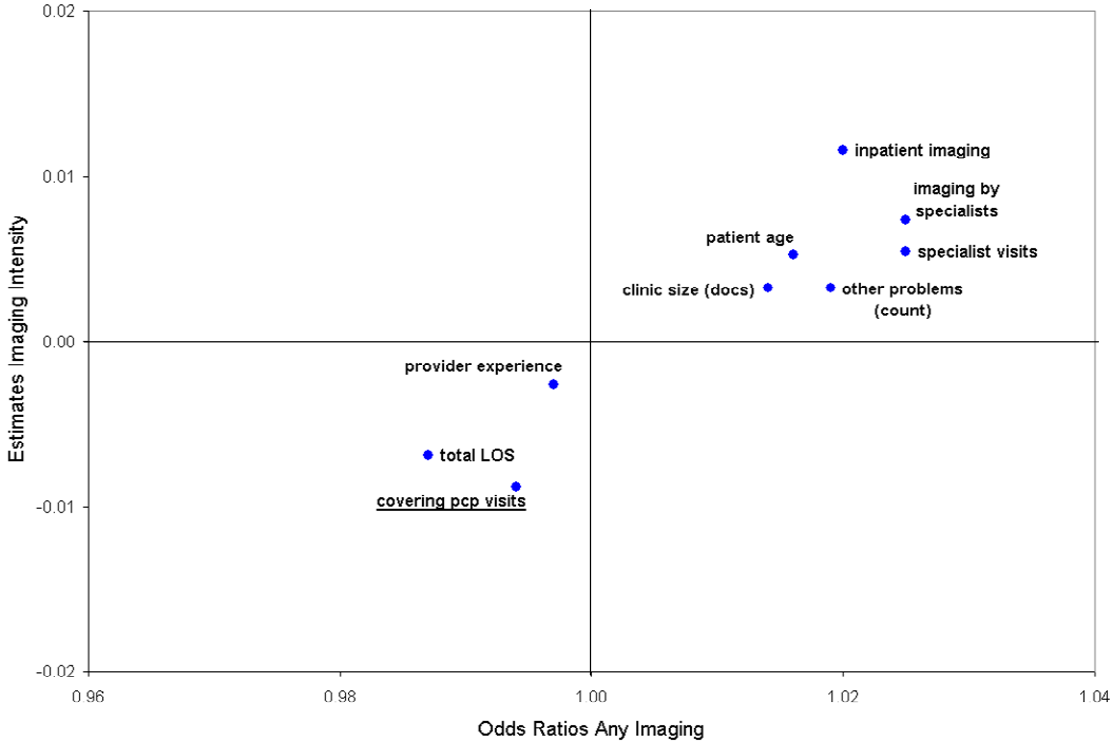

Figure 7-5. Comparison of significant variables for any imaging use and imaging intensity (small effect sizes). Unless underlined, all variable/levels were significant for both any imaging use and imaging intensity. The summed RVU of visits to covering primary care doctors (underlined) was not significant for any imaging use.

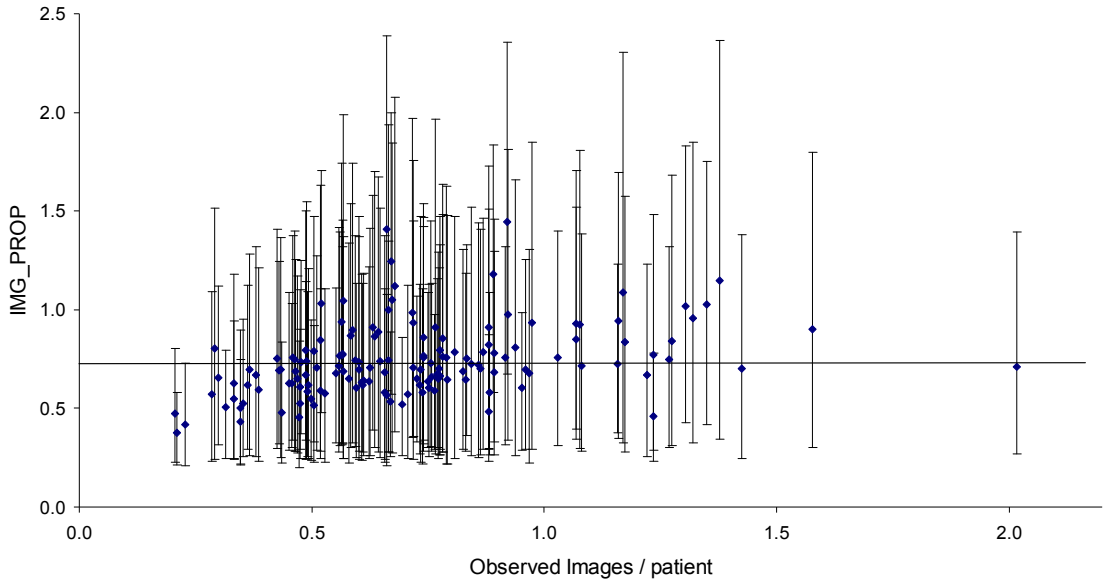

Figure 7-6. Imaging propensity score distributions by provider. Each provider's (N=148) mean imaging propensity (diamonds) along with 10th and 90th percentiles (error bars) are plotted on the Y axis against the observed mean number of images per patient for that provider (X axis). The grand mean of imaging propensity (0.73) is indicated by the horizontal line.

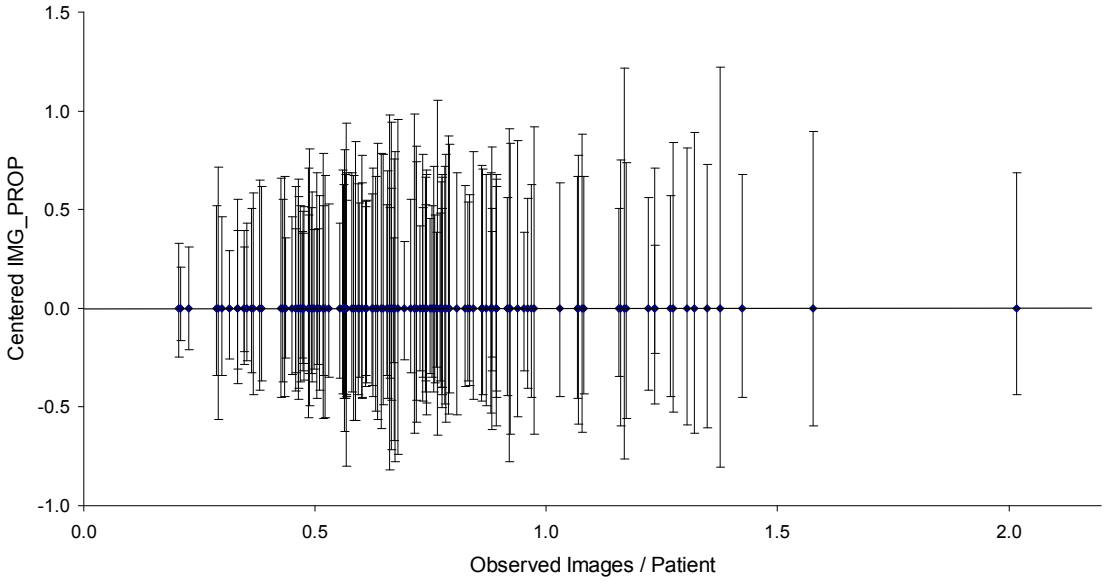

Figure 7-7. Centered imaging propensity score distributions by provider. Each provider's (N=148) centered imaging propensity (diamonds) along with 10th and 90th percentiles (error bars) are plotted on the Y axis against the observed mean number of images per patient for that provider (X asis). The overall centered mean of imaging propensity (zero) is indicated by the horizontal line.

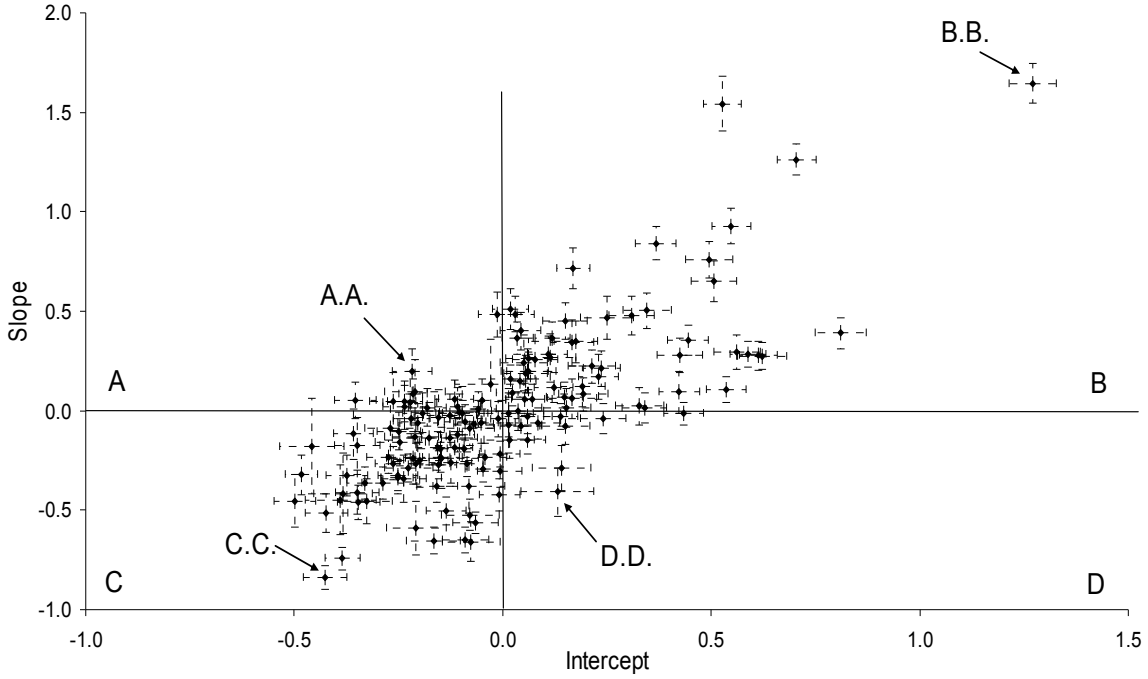

Figure 7-8. Plot of intercept and slopes for all 148 providers obtained from multi-level model of imaging utilization. Error bars represent ± standard error for each provider's intercept (horizontal) and slope (vertical).

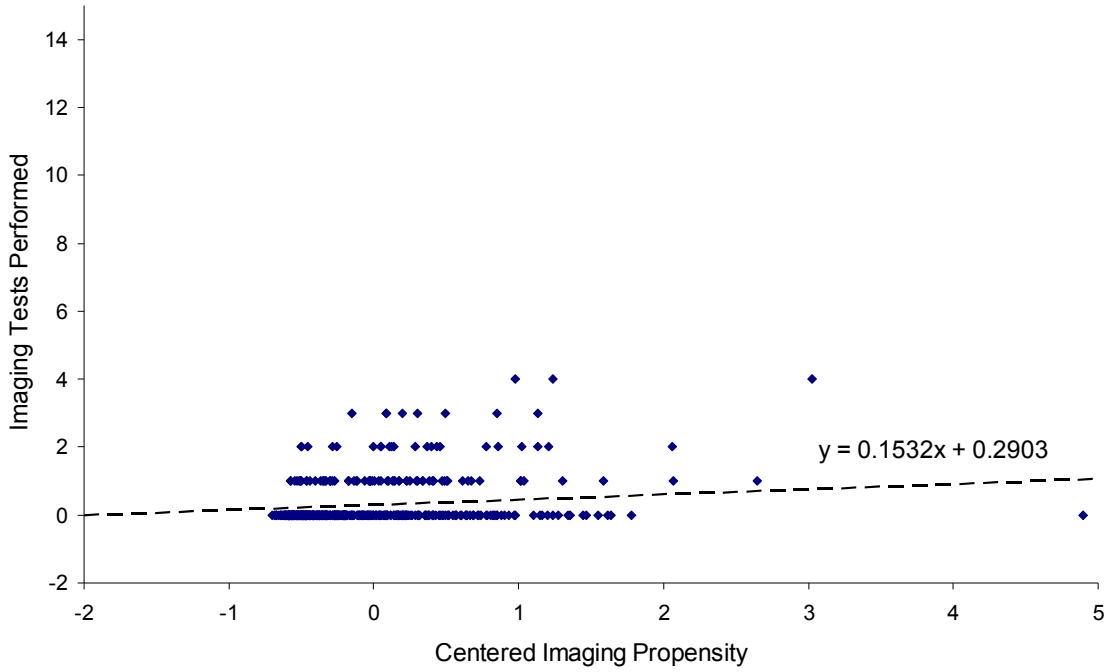

Figure 7-9. Imaging utilization versus centered imaging propensity for a low utilizing doctor (C.C. in Figure 7-8). Each diamond represents a single patient and the dashed line is the linear equation defined by the multi-level model intercept and slope for provider C.C.

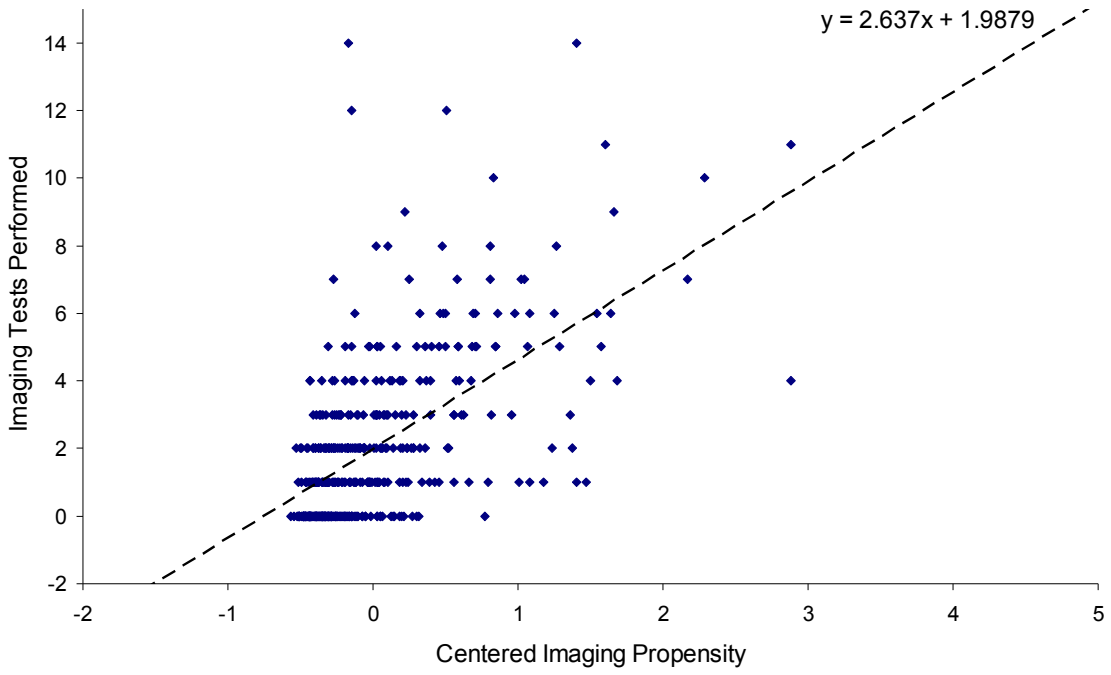

Figure 7-10. Imaging utilization versus centered imaging propensity a high utilizing doctor (B.B. in Figure 7-8). Each diamond represents a single patient and the dashed line is the linear equation defined by the multi-level model intercept and slope for provider B.B.

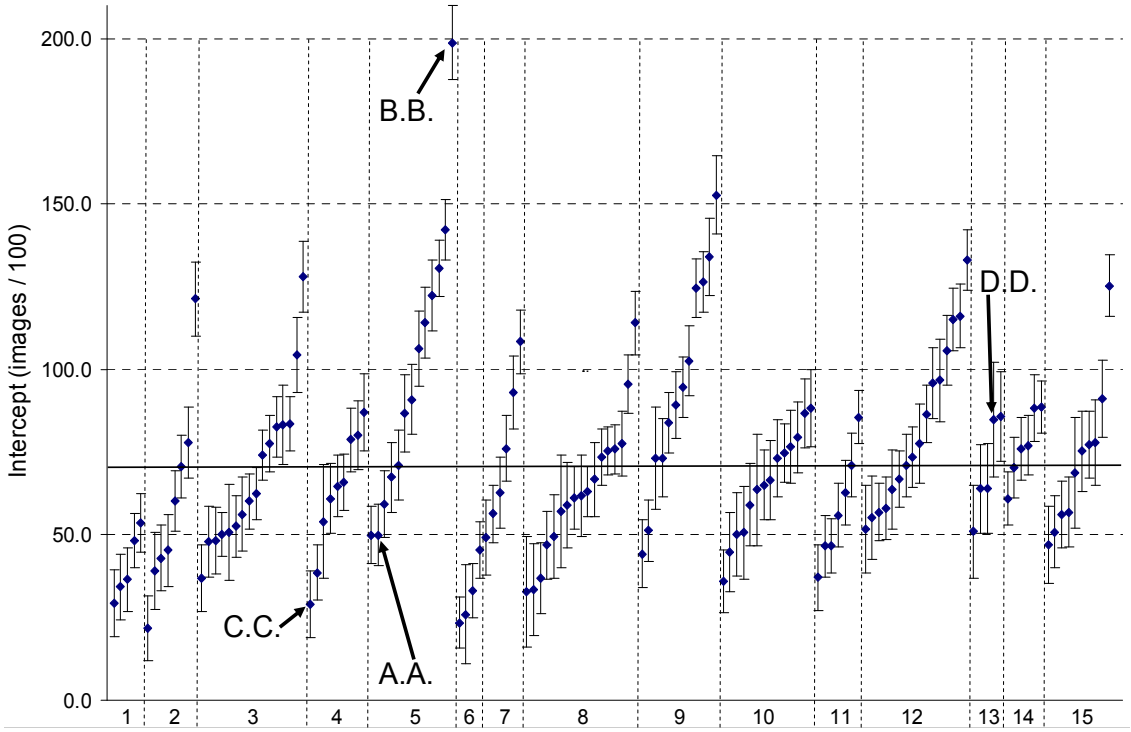

Figure 7-11. Provider means sorted by ascending order within each site. Site (clinic) numbers shown along the bottom. Each provider's intercept is scaled by adding the model coefficient for each provider to the fixed intercept (0.7171) and multiplied by 100 to represent adjusted images per 100 patients over the two year study interval. The solid horizontal line is at 71.7 images per 100 patients which is the grand mean number of images per 100 patients. Error bars are 95% confidence intervals scaled up the same way. The individual providers (A.A. – D.D.) as discussed in the text are labeled).

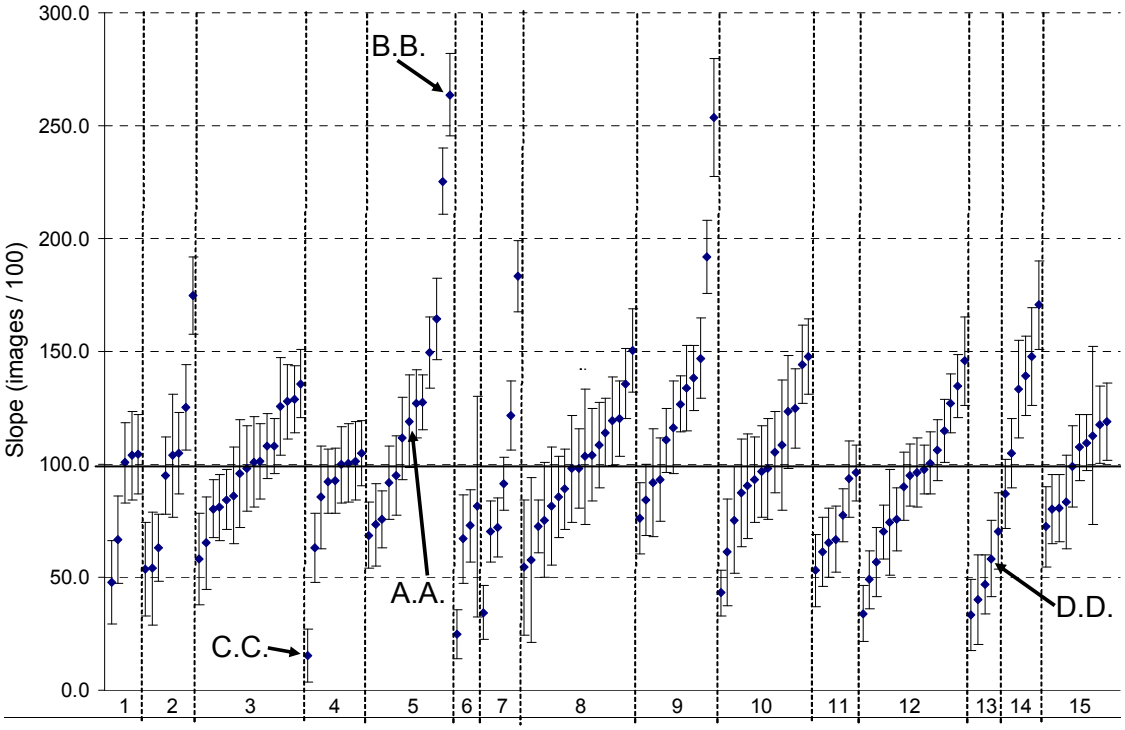

Figure 7-12. Provider slopes sorted by ascending order within each site. Site (clinic) numbers shown along the bottom. Each provider's sloped is scaled by adding the model coefficient for each provider to the fixed imaging propensity effect (0.9919) and multiplied by 100 to represent adjusted images per 100 patients over the two year study interval. Error bars are 95% confidence intervals scaled up the same way. The individual providers (A.A. – D.D.) as discussed in the text are labeled).

CHAPTER 8
DISCUSSION

Summary of Key Results

Utilization of outpatient diagnostic imaging in a cohort of 85,277 patients was evaluated over a two year period extending through June of 2009. The adult patients in this study were cared for in a stable 'medical home' as defined by regular visits to a primary care doctor practicing in one of 15 clinics. The institutional setting is an academic health center located in Boston. In general, the study revealed that older female patients had more imaging as did those who had many medical problems listed in the clinical record system. Also, patients who visited doctors, were admitted to the hospital, or seen in the emergency room more frequently had more imaging. Doctor factors associated with a greater tendency to order imaging tests were less experience, female gender, and having a medium size practice (500-1000 patients). A special statistical technique (hierarchical modeling) that accounts for all patient factors allowed creation of 'profiles' scoring each of the 148 doctors on their general tendency to order imaging tests on the 'average' patient and how many more tests were ordered on patients with greater comparative 'need' for diagnostic imaging.

On a per patient basis, the amount of imaging ordered by their own (linked / loyal) primary care doctor ranged from 0-15 examinations, with average of 0.7146 (60,938 / 85,277). This translates into 35.7 images per 100 patient years. For comparison, during the same period, the cohort had 50.8 outpatient images per 100 patient years ordered by specialists caring for them and 37.6 images per 100 patient years performed while they were in the emergency department or hospitalized. This study concentrated exclusively on variability in and factors contributing to the outpatient imaging ordered by

the linked (loyal) primary care doctor, which accounts for about 27.5% of the total imaging received by the entire patient cohort.

With the patient experience over 2 years as the unit of analysis and the count of primary care ordered outpatient images as the outcome, fitting a simple Poisson distribution yields mean (λ) of 0.715 (CI: 0.707-0.722). The outcome distribution also fits quite well to a Zero-Inflated Poisson (ZIP) distribution with mean (λ) of 1.492 (CI: 0.389-0.411) and zero-inflation parameter (ϕ) of 0.521 (CI: 0.516-0.526). This (ZIP) distribution was used for model based creation of patient level imaging propensity expected number of images using available risk adjustment variables.

An alternate way to model the same phenomenon is in two stages; any imaging versus none (logistic process) followed by imaging intensity (Poisson process) for patients with at least one imaging test. For the logistic process, the overall 'success' rate (patient had some imaging) was 37.13% (31,660 / 85,277). The Poisson mean (λ) for the non-zero imaging (N=31,660 patients) was 1.925 (CI: 1.910-1.940). This two stage approach was used to test joint effect of 28 patient, 6 provider, and 1 clinic variable(s) on any imaging utilization (logistic regression) followed by imaging intensity (Poisson regression).

Patient demographic and clinical factors significantly associated with a greater likelihood of any imaging *and* greater imaging intensity included: increasing patient age, female sex, Hispanic race (compared with white), and more clinical problems. Patient level clinical activity variables associated with greater likelihood of any imaging *and* greater imaging intensity included: office visits to the linked (loyal) primary care doctor, office visits to specialists, inpatient imaging tests, outpatient imaging ordered by other

(covering) primary care doctors, and outpatient imaging ordered by specialists. Provider and clinic level factors associated with a greater likelihood of any imaging and greater imaging intensity included: an extra degree (after M.D.) held by the patient's linked (loyal) primary care doctor, patient's linked (loyal) doctor was female, mid-level practice size (500-799 patients), and increasing size of the clinic (number of practicing doctors).

Patient level factors associated with decreased likelihood of any imaging and decreased imaging intensity included: insurance with Medicare (compared to self-pay), diabetes, and hypertension. Patients with longer total length of stay in the hospital actually had lower likelihood of any imaging and lower imaging intensity. As the primary care provider's experience increased, both likelihood and intensity of imaging for their linked (loyal) patients decreased. The provider's place of M.D. training (foreign medical graduate=FMG) had a discordant effect on imaging utilization. When the doctor had FMG status of 'yes', they tended to be more likely to order at least some imaging on their patients but the amount of imaging tests was less than American trained (FMG='no') counterparts.

One factor that did not have any significant effect during multivariable modeling of imaging utilization (any imaging or imaging intensity) was the amount of outpatient prescription medications (in 4 ordinal categories) each patient was taking. This is notable because there was a strong bivariate relationship between this medication variable and the outcome (outpatient imaging ordered by primary care doctor). This will be discussed below.

When all patient level factors were combined in a single ZIP model, the predictions for each patient were used as an (expected) imaging propensity score for subsequent

multi-level hierarchical modeling. This imaging 'risk adjustment' model had R-Squared of 0.17 at the patient level. Adding the provider's unique (anonymous) identity as a predictor in a two level hierarchical model brought the full model R-Squared up to 0.24. Since the patient level imaging propensity scores were centered on each provider's mean, the provider intercept can be interpreted as the average imaging utilization for that doctor (all else equal). The plausible range of these intercepts was 0.151 to 1.283 images per patient over two years with mean of 0.7171. This translates into 35.9 images per 100 patient years with plausible range of 7.6 to 64.2. At the same time, the individual provider slopes from the hierarchical model can be interpreted as the extent to which each doctor responds to a unit increase in patient imaging propensity (which also ranged from 0 to 15 with a single higher score of ~20). The mean was 0.9919 and the plausible range was 0.216 to 1.768. Scaling these up to images per 100 patient years gives mean of ~50 and plausible range of 10.8 to 88.4 for every unit increase in patient imaging 'need'. These estimates of provider imaging utilization parameters (average/intercept and slope) are quite precise with calculated reliability of 0.965 and 0.939 respectively.

Discussion of Key Results

Perhaps the most vexing and interesting question arising from this study has to do with the fact that no more than 25% of the variation in the number of primary care imaging tests per patient is explainable using a quite robust and complete set of patient, doctor, and clinic variables. This holds true for even the most complete models tried. For example, the ZIP model with all 28 patient level variables that produced the imaging propensity scores for multi-level modeling can be modified to include *all* variation due to providers and clinics by placing the unique identity of each patient's doctor and clinic

into the model (as class variables). The R-Squared of this ‘fullest’ model is still just over 0.24 and about 30% of the explanatory ‘power’ comes from knowing who the doctors and clinics are. Further, the empirically determined variation in the 148 providers tendency to order imaging on the average patient (intercepts and slopes from multi-level modeling) is quite substantial. This conundrum goes to the very heart of philosophical considerations concerning causes and consequences of variations in medical resource utilization. These questions have been posed occasionally in the health services research literature but remain unanswered (Cain and Diehr 1992, Diehr *et al.* 1990).

With the preceding in mind, consider that the practice setting (MGH) is among the most sophisticated and consistent with respect to the processes of outpatient imaging ordering, scheduling, and provider feedback. As mentioned in the Chapter 5, virtually all outpatient imaging was ordered and scheduled via a web-based radiology order entry (ROE) system. During the entire study period, the ROE system had fully functional and complete real-time appropriateness decision support (DS) feedback for all CT, MRI, and nuclear medicine tests. Additionally, the primary care doctors included in this study were all given periodic (bi-yearly) feedback about their utilization of outpatient imaging compared with peers. It can be persuasively argued that the practice examined herein is ‘as good as it gets’ with respect to outpatient imaging decision support and utilization management. This implies that the amount of variability in primary care outpatient imaging utilization accruing to doctors (~30%) is a lower bound. Therefore, if similar studies were to be conducted elsewhere, the absolute amount of variability between doctors would be greater and the fraction of total variation also larger.

Limitations of Study

In terms of generalizability, the primary care practice being studied is not widely representative of other private, or even other academic settings. As mentioned above, near complete use of electronic order entry with imaging-specific decision support coupled with active utilization management means that most other practice settings and locales will differ on several axes. The absolute number (and distribution over modality) of images obtained per patient in other primary care settings may be substantially higher or lower. Other primary care doctors/groups may actually obtain advanced imaging less often without easy access to electronic ordering, scheduling, and clinical decision support. For, example they may more often refer complex patients to specialists and defer imaging to them resulting in lower apparent utilization by the primary care provider(s). On the other hand, without 'barrier' effects of a formal order entry system and decision support, which sometimes recommends against imaging, overall utilization could be much higher. In either event, variability of utilization between patients (after risk adjustment) and providers will almost certainly be greater in other settings. This would manifest in a lower fraction of overall explained variation (25% in this setting), less effective risk adjustment at the patient level, and a greater fraction of variation attributable to providers. These observations should not discourage others from using risk adjusted benchmarking of imaging utilization. Quite to the contrary, such provider profiles under conditions of greater variation in imaging utilization will have potentially greater impact.

The main study outcome, outpatient imaging utilization, was quantified by counting imaging tests. Clearly, not all imaging tests are equal and some cost more others. For utilization management efforts that seek to understand and control expenditures,

resource use (cost) of each imaging procedure is important. The relative value unit (RVU) of each imaging procedure is the obvious choice as a proxy for cost and is in fact often used to calculate reimbursement. Thus, the summed RVU of outpatient imaging tests performed on each patient has a potential advantage over simple counts if cost were to be the main focus of analysis. For example, an 'old school' doctor who ordered chest X-Ray on most patients having respiratory symptoms might seem to have the same level of utilization (by simple counting of procedures) as a doctor that ordered chest CT scans much more often. In comparing relative contribution of primary care and specialist ordered tests to outpatient imaging expenditures, summed RVU (as opposed to simple counts) would account for differences in the type of tests (modalities) that get ordered. Also, the summed RVU as outcome approach would allow model based prediction/speculation about potential cost savings that could be realized by reducing provider variation and/or curtailing utilization among the 'high outliers'. This study only examined outpatient imaging ordered by each patient's linked primary care doctor which accounts for less than half (~40%) of all outpatient imaging. The vast majority of the rest is ordered by the (sometimes many) specialists caring for the same patient. However, the relationship between primary care and specialist ordered imaging was partly addressed by including the amount specialist ordered imaging as a patient level predictor.

The assumption about distribution of errors in modeling counts of imaging was that they were Poisson or ZIP. In nature, true Poisson data generating processes have identical probability of events during time t+1 independent of the cumulative number of events through time t. It can be argued that in actual patient care, this independence

assumption may not hold because ‘sicker’ patients tend to have more imaging tests in the future. At the same time, patients who have been ill may already had had more imaging tests in the past than otherwise healthy peers such that prior imaging is a ‘marker’ for having been sick. However, the distribution of imaging test counts over patients has an empirical shape that is quite well fitted by Poisson or ZIP distributions. At the same time, one of the main reasons for selecting non normal error distribution for count modeling is to avoid deflation of standard errors of estimates and resulting mistakes in hypothesis tests about them. Thus, even if the data generating phenomenon is not a perfect ‘natural’ Poisson or ZIP process, these distributions may still be most suitable for error fitting.

The multivariable modeling of any imaging use (logistic) and imaging intensity (Poisson on non-zero observations) did not account for the nested structure of the data so that the doctor and clinic factors were repeated over all patients in each respective unit. This may result in biased estimation of the effect size and somewhat lower standard errors for these coefficients. The subsequent hierarchical modeling helps to address this shortcoming. Despite selecting one from each group of highly correlated variables, there may be additional problems with multi-colinearity as evidenced by the behavior of the variable measuring patient’s outpatient medications (significant at bivariate analysis but not significant when analyzed jointly with other variables). Also, additional work will better characterize the residuals from the logistic, Poisson, ZIP, and multi-level models used at various stages of the analysis. Alternate error distribution assumptions for the count of imaging tests per patient might also be worthy of exploration, including Negative Binomial.

There were 10,396 (12.2%) patients that had no recorded visits to the linked (loyal) provider during the two year study period. These were retained in the analytic data set. This raises the concern that the linked (loyal) provider visit RVU variable (which was strongly associated with the number of imaging tests ordered by that doctor) was confounded in some way or that the loyalty attribution methodology was flawed. By definition of the (2008) loyalty cohort, all patients had at least one visit to their linked provider from 2006-2008 and this was confirmed. All 10,396 patients had at least one visit to the linked provider occurring between January 1, 2006 (start of the loyalty cohort definition period) and July 1, 2007 (start of the study period). The patients with no visits during the study period were distributed across 143 of the providers. That is, only 5 providers saw all their loyal patients at least once during the study. A by provider distribution of the percent patients with no visit during the study period had mean=11.8, median=11.3, and standard deviation=8.25 in a nearly normal distribution (skew=1.11) of loyal patients who did not visit them during the study period. This is reassuring in that it would seem to reflect actual practice variation rather than a substantial data integrity problem (e.g., provider identifier mismatch).

There were 804 patients excluded from the analytic data set because they were cared for by one of 26 physicians that had less than 100 loyal patients. In general, these were doctors who had mixed practices, worked part time, or left practice during the study period. For example, an endocrinologist (cardiologist, gastroenterologist) that worked out of a primary care clinic might have a few patients identified as ‘loyal’ to them by the Atlas methodology. Alternatively, a doctor with administrative, teaching, or research commitments taking up most of their time might attend in one of the primary

care clinics a few days per month. Of the 804 excluded patients 593 (74%) had no images ordered by the linked (loyal) provider, 764 (95%) had two or less, and 40 (5%) had from 3-7 images. Exclusion of patients (and doctors) in these small and/or mixed practices may reduce generalizability. However, the specific aim of this research was to evaluate imaging utilization by actively practicing primary care doctors and including this small number of patients and their doctors might have biased the results without adding any additional useful information.

The hierarchical modeling was only carried out with two levels (patients and doctors) which discounts the effect of having patients nested within doctors that are in turn nested in clinics. Perhaps a single grand three level (patient, doctor, and clinic) model that incorporates all relevant predictors individually and handles error distribution robustly (e.g. ZIP) might provide greater insight into the phenomenon and produce superior estimates of parameters. This would be a very complex undertaking and might well involve a full ZIP specification (primary and zero-model) at each of 3 hierarchical levels (patient, doctor, and clinic).

Policy Implications

Two overarching policy concerns attach to outpatient diagnostic imaging; substantial and rapidly rising costs as well as increasing population radiation burden from medical imaging (especially due to CT scans) with attendant risk of cancer induction. In addition, the situation of primary care doctors deciding between test (imaging), treat, observe, or refer when faced with their patient's varying clinical presentations is a classic paradigm of medical decision making. Any attempts to curtail imaging utilization growth in general or to target 'high users' for remediation or sanction must be informed by proper modeling of drivers of variation at patient, provider, and

perhaps higher levels. The empiric findings of this study will help to understand the manner and extent to which risk adjustment and variation analysis methods can help with outpatient imaging utilization management. For example, there is a growing body of literature reflecting a vigorous debate about the reliability, utility, and fairness of provider 'efficiency' profiles promulgated by payers (Adams *et al.* 2010). Unlike typical 'observed / expected' metrics, the multi-level modeling described herein, directly produces highly reliable (>95%) and much more meaningful provider level measures about average utilization (intercept) and response to patient clinical need variables (slope). Further, each provider intercept and slope has its own standard error which allows much more meaningful comparison with individual peers and the overall average utilization.

As stated above, the clinical leadership for the large group practice serving as the setting for this study has engaged in quite robust and longstanding utilization management efforts specific to outpatient imaging utilization and appropriateness. Also, the doctors studied were all salaried employees of the group practice and none had any financial incentives (or disincentives) associated with diagnostic imaging. For example, there was almost no growth in the use of CT scans by this practice for 4 years which include the period of study (Sistrom *et al.* 2009). At the same time, double digit rates of growth in CT volumes have occurred in many other settings in the U.S. Thus, the average yearly number of imaging tests ordered by these primary care doctors (~36 per 100 patients) would seem to be a lower bound estimate of what occurs nationally. At the same time, there was substantial variation between doctors in their average use of imaging (intercept from multi-level model) and their response to patient propensity (clinical need) for imaging (slopes from multi-level analysis).

There are policy implications from this study that relate to high technology in medical care more generally. Advanced imaging is a ‘poster child’ for a broad array of diagnostic and therapeutic interventions made possible by advances in basic sciences, engineering, and informatics. These devices and techniques are attractive and compelling to health care providers, patients, and lay public such that hospitals now actively compete to obtain and advertise extensively about the ‘latest and greatest’ advances. Further, the regulation of medical devices by the Food and Drug Administration and other agencies is much less stringent than for drugs. Specifically, there is little or no requirement that developers and vendors demonstrate clinical effectiveness; only safety and functionality. Comparative effectiveness evaluations to determine appropriateness of various devices and technologies for different clinical purposes will be needed to guide reimbursement determinations. These have two separate stages: first, whether or not to allow claims for a new device/technology at all, and second, what clinical situations warrant reimbursement on a case by case basis. As with imaging, there are many contextual factors surrounding utilization of high technology medical interventions that operate in concert (or opposition) with clinical need (appropriateness).

Contribution to Literature

The use of the MGH/Atlas loyalty cohort methodology provides a unique population of patients and doctors participating in a stable ‘medical home’ type of primary care practice. As described above, the robust medical and imaging informatics infrastructure at MGH provides an optimal situation for standardizing the appropriateness and intensity of outpatient imaging utilization. Thus the practice under

observation is quite likely at or near optimum with respect to variation between primary care providers with respect to diagnostic imaging. This same rich informatics environment also means that the available empirical data about patients, doctors, and the clinical activity they engage in (including but not limited to imaging tests) is unparalleled in fidelity and completeness. The estimates of the effect size and direction of many patient and provider level factors on imaging utilization should be useful in themselves.

Despite increasing popularity and application of hierarchical techniques to health outcomes, cost analyses, and risk-adjustment, this is the first study in which hierarchical modeling has been used to study outpatient imaging utilization in primary care. Preparatory risk-adjustment for 'imaging propensity' at the patient level using ZIP modeling is also unique. The combination of these methods yields important and interesting insights into how doctors differ in both their general tendency to use imaging on the 'average' patient but how they respond to changing 'need' for imaging in their own panel of patients.

Future Research

The current data source includes individual records for every outpatient visit with CPT code indicating type and intensity of visit, ICD-9 codes for the visit reason(s), patient ID, rendering provider ID, and date of service. The CPT codes were already used to create summed RVU of visits by provider type for the completed study. Combining these, it should be possible to use the visit data in much more robust ways. For example, the date, provider, and patient information common to both visit records and outpatient imaging events can be submitted to 'attribution logic' which matches imaging tests to visits. A simple set of rules serves to do this. Also, it is important to

have at least two contiguous years worth of visits and three years of imaging data on the same population (which is available for this cohort).

This ‘visit-based’ method has been validated on a separate set of neurology outpatient visits and associated imaging tests with successful attribution of more than 90% of imaging tests to a visit. The analysis becomes a visit-based rather than panel-based and the ‘measure’ becomes images per visit rather than images per patient year. By modeling visits and grouping by provider, comparison between providers concerning their relative tendency to order imaging can be performed. Using the same set of patients and primary care doctors over the same time frame would allow comparing visit-based and panel-based imaging utilization profiles to see if the visit-based method gives similar results (e.g., ranks providers in the same order in terms of imaging utilization intensity). If a visit-based method is acceptable, it can be generalized to data sets that are less granular and robust (e.g., Medicare claims). The other advantage of visit-based provider profiles is that they will work with specialists who are much less likely--than primary care doctors--to have a stable ‘panel’ of patients.

Another interesting set of questions arising from the current data source has to do with the relationship between outpatient imaging ordered by the patient’s linked (loyal) primary care doctor and other doctors, mostly specialists. Specifically, in this cohort of patients the majority (about 60%) of all outpatient imaging was ordered by specialists. As described above, primary care doctors faced with clinical uncertainty have a limited set of options: observe, treat, (imaging) test, or refer to specialist. The current study lumps three of the choices; observe, treat, or refer into a ‘no imaging’ category and evaluates factors relating to the single alternate choice: order an imaging test. Parallel

and/or simultaneous modeling (perhaps with multivariate techniques) of the imaging performed by specialists could add an additional dimension to our characterization of primary care doctor behavior. It may turn out that some providers who seem to be ‘conservative’ with respect to imaging are actually ‘liberal’ in terms of referring to specialists and this behavior may costlier overall compared with providers who order more of their own imaging tests.

LIST OF REFERENCES

Adams, J. L., A. Mehrotra, J. W. Thomas, et al. 2010. Physician Cost Profiling—Reliability and Risk of Misclassification. *New England Journal of Medicine 362:1014–1021.*

Allan, G. M., and J. Lexchin. 2008. Physician Awareness of Diagnostic and Nondrug Therapeutic Costs: A Systematic Review. *International Journal of Technology Assessment in Health Care* 24:158–165.

Andersen, R. M. 1995. Revisiting the Behavioral Model and Access to Medical Care: Does it Matter? *Journal of Health and Social Behavior* 36:1–10.

Anthony, D. L., M. B. Herndon, P. M. Gallager, et al. 2009. How Much Do Patients' Preferences Contribute to Resource Use? *Health Affairs (Millwood)* 28:864–873.

Atlas S. J., Y. Chang, T. A. Lasko, et al. 2006. Is this "My" Patient? Development and Validation of a Predictive Model to Link Patients to Primary Care Providers. *Journal of General Internal Medicine* 21:973–978.

Atlas, S. J., R. W. Grant, T. G. Ferris, et al. 2009. Patient-Physician Connectedness and Quality of Primary Care. *Annals of Internal Medicine* 150: 325–335.

Axt-Adam, P., J. C. van der Wouden, and E. van der Does. 1993. Influencing Behavior of Physicians Ordering Laboratory Tests: A Literature Study. *Medical Care* 31:784–794.

Baicker, K., E. S. Fisher, and A. Chandra. 2007. Malpractice Liability Costs and the Practice of Medicine in the Medicare Program. *Health Affairs (Millwood)* 26:841–852.

Bates, D. W., G. J. Kuperman, A. Jha, et al. 1997. Does the Computerized Display of Charges affect Inpatient Ancillary Test Utilization? *Archives of Internal Medicine* 157:2501–2508.

Bernardy, M., C. G. Ullrich, J. V. Rawson, et al. 2009. Strategies for Managing Imaging Utilization. *Journal of the American College of Radiology* 6:844–850.

Berrington de Gonzalez, A., M. Mahesh, K. P. Kim, et al. 2009. Projected Cancer Risks from Computed Tomographic Scans Performed in the United States in 2007. *Archives of Internal Medicine* 169:2071–2077.

Bertolini, G., R. D'Amico, D. Nardi, et al. 2000. One Model, Several Results: The Paradox of the Hosmer-Lemeshow Goodness-of-Fit Test for the Logistic Regression Model. *Journal of Epidemiology and Biostatistics* 5:251–253.

Bhargavan, M., and J. H. Sunshine. 2005. Utilization and Radiology Services in the United States: Levels and Trends in Modalities, Regions, and Populations. *Radiology* 234:824–832.

Blachar, A., S. Tal, A. Mandel, et al. 2006. Preauthorization of CT and MRI Examinations: Assessment of a Managed Care Preauthorization Program based on the ACR Appropriateness Criteria and the Royal College of Radiology Guideliness. *Journal of the American College of Radiology* 3:851–859.

Blumberg, M. S. 1987. Inter-area Variations in Age-adjusted Health Status. *Medical Care* 25:340–353.

Bodenheimer, T., and A. Fernandez. 2005. High and Rising Health Care Costs. Part 4: Can Costs be Controlled While Preserving Quality? *Annals of Internal Medicine* 143:26–31.

Boutwell, R. C. and J. B. Mitchell. 1993. Diffusion of New Technologies in the Treatment of the Medicare Population. Implications for Patient Access and Program Expenditures. *International Journal of Technology Assessment in Health Care* 9:62–75.

Brant-Zawadzki, M. 1994. Cost Containment Drives Utilization Management. *Diagnostic Imaging (San Francisco)* 16:37–43.

Brenner, D. J., and E. J. Hall. 2007. Computerized Tomography—An Increasing Source of Radiation Exposure. *The New England Journal of Medicine* 357:2277–2284.

Britt, H., A Bhasale, D. A. Miles, et al. 1996. The Sex of the General Practitioner: A Comparison of Characteristics, Patients, and Medical Conditions Managed. *Managed Care* 34:403–415.

Brook, R. H., and K. N. Lohr. 1985. Efficacy, Effectiveness, Variations, and Quality. Boundary-crossing Research. *Medical Care* 23:710–722.

Bugter-Maessen, A. M., R.A. Winkens, R. P. Grol, et al. 1996. Factors Predicting Differences among General Practitioner in Test Ordering Behaviour and in the Response to Feedback on Test Requests. *Family Practice* 13:254–258.

Burkhardt, J. H., and J. H. Sunshine. 1996. Utilization of Radiologic Services in Different Payment Systems and Patient Populations. *Radiology* 200:201–207.

Cain, K. C., and P. Diehr. 1992. Testing the Null Hypothesis in Small Area Analysis. *Health Services Research* 27:267–294.

Casparie, A. F. 1996. The Ambiguous Relationship between Practice Variation and Appropriateness of Care: An Agenda for Further Research. *Health Policy* 35:247–265.

Chang, W. and S. B. McCracken. 1996. Applying Case Mix Adjustment in Profiling Primary Care Physician Performance. *Journal of Health Care Finance* 22:1–9.

Chassin, M. R. 1993. Explaining Geographic Variation. The Enthusiasm Hypothesis. *Medical Care* 31:YS37–YS44.

Chassin, M. R., J. Kosecoff, R. E. Park, et al. 1987. Does Inappropriate Use Explain Geographic Variations in the Use of Health Care Services? A Study of Three Procedures. *Journal of the American Medical Association* 258:2533–2537.

Childs, A. W., and E. D. Hunter. 1972. Non-medical Factors Influencing Use of Diagnostic X-ray by Physicians. *Medical Care* 10:323–335.

Cohen, B. J. 1996. Is Expected Utility Theory Normative for Medical Decision Making? *Medical Decision Making* 16:1–6.

Couchman, G. R., S. N. Forjuoh, M. H. Rajab, et al. 2004. Nonclinical Factors Associated with Primary Care Physicians' Ordering Patterns of Magnetic Resonance Imaging/Computed Tomography for Headache. *Academic Radiology* 11:735–740.

Couchman, G. R., S. N. Forjuoh, M. D. Reis, et al. 2005. Variation in MRI/CT Utilization among Family Physicians and General Internists in a Multi-specialty Group Practice. *Medical Science Monitor* 11:MT19–MT25.

Cummings, K. M., K. B. Frisof, M. J. Long, et al. 1982. The Effects of Price Information on Physicians' Test-ordering Behavior. Ordering of Diagnostic Tests. *Medical Care* 20:293–301.

Davis, P., B. Gribben, R. Lay-Yee, et al. 2002. How Much Variation in Clinical Activity is There Between General Practitioners? A Multi-Level Analysis of Decision-Making in Primary Care. *Journal of Health Services Research & Policy* 7:202–208.

Davis, P., B. Gribben, A. Scott, et al. 2000a. Do Physician Practice Styles Persist Over Time? Continuities in Patterns of Clinical Decision-Making Among General Practitioners. *Journal of Health Services Research & Policy* 5:200–207.

Davis, P., B. Gribben, A. Scott, et al. 2000b. The "Supply Hypothesis" and Medical Practice Variation in Primary Care: Testing Economic and Clinical Models of Inter-Practitioner Variation. *Social Science & Medicine* 50:407–418.

DeKay, M. L. and D. A. Asch 1998. Is the Defensive Use of Diagnostic Tests Good for Patients, or Bad? *Medical Decision Making* 18:19–28.

Deyo, R. A. 2002. Cascade Effects of Medical Technology. *Annual Review of Public Health* 23:23–44.

Diehr, P., K. Cain, F. Connell, et al. 1990. What is too Much Variation? The Null Hypothesis in Small-Area Analysis. *Health Services Research* 24:741–771.

Eddy, D. 2007. Reflections on Science, Judgment, and Value in Evidence-Based Decision Making: A Conversation with David Eddy by Sean R. Tunis. *Health Affairs (Millwood)* 26:w500–w515.

Eddy, D. M. 1984. Variations in Physician Practice: The Role of Uncertainty. *Health Affairs (Millwood)* 3:74–89.

Eddy, D. M. 1990. Clinical Decision Making: From Theory to Practice. Guidelines for Policy Statements: The Explicit Approach. *Journal of the American Medical Association* 263:2239–40, 2243.

Eeckhoudt, L. 1996. Expected Utility Theory--Is it Normative or Simply "Practical"? *Medical Decision Making* 16:12–13.

Eisenberg, J. M. 1985. Physician Utilization. The State of Research about Physicians' Practice Patterns. *Medical Care* 23:461–483.

Eisenberg, J. M. 1986a. The Physician as Self-Fulfilling Practitioner. in *Doctor's Decisions and the Cost of Medical Care*, Health Administration Press Perspectives, Ann Arbor, MI, pp. 29–56.

Eisenberg, J. M. 1986b. The Physician as the Patient's Agent. in *Doctor's Decisions and the Cost of Medical Care*, Health Administration Press Perspectives, Ann Arbor, MI, pp. 57–78.

Eisenberg, J. M., and D. Nicklin. 1981. Use of Diagnostic Services by Physicians in Community Practice. *Medical Care* 19:297–309.

Epstein, A. M., and B. J. McNeil. 1985a. Physician Characteristics and Organizational Factors Influencing Use of Ambulatory Tests. *Medical Decision Making* 5:401–415.

Epstein, A. M., and B. J. McNeil. 1985b. The Effects of Patient Characteristics on Ambulatory Test Ordering. *Social Science & Medicine* 21:1071–1075.

Epstein, A. M., and B. J. McNeil. 1986. Relationship of Beliefs and Behavior in Test Ordering. *The American Journal of Medicine* 80:865–870.

Epstein, A. M., and B. J. McNeil. 1987. Variations in Ambulatory Test Use: What Do They Mean? *Medical Clinics of North America* 71:705–717.

Fazel, R., H. M. Krumholz, Y. Wang, et al. 2009. Exposure to Low-Dose Ionizing Radiation from Medical Imaging Procedures. *The New England Journal of Medicine* 361:849–857.

Fenn, P., A. Gray, and N. Rickman. 2007. Liability, Insurance and Medical Practice. *Journal of Health Economics* 26:1057–1070.

Fertig, A., M. Roland, H. King, et al. 1993. Understanding Variation in Rates of Referral Among General Practitioners: Are Inappropriate Referrals Important and Would Guidelines Help to Reduce Rates? *British Medical Journal* 307:1467–1470.

Fisher, E. S., D. E. Wennberg, T. A. Stukel, et al. 2003a. The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care. *Annals of Internal Medicine* 138:273–287.

Fisher, E. S., D. E. Wennberg, T. A. Stukel, et al. 2003b. The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care. *Annals of Internal Medicine* 138:288–298.

Folland, S., and M. Stano. 1989. Sources of Small Area Variations in the Use of Medical Care. *Journal of Health Economics* 8:85–107.

Folland, S., and M. Stano. 1990. Small Area Variations: A Critical Review of Propositions, Methods, and Evidence. *Medical Care Review* 47:419–465.

Franks, P., and K. Fiscella. 2002. Effect of Patient Socioeconomic Status on Physician Profiles for Prevention, Disease Management, and Diagnostic Testing Costs. *Medical Care* 40:717–724.

Franks, P., C. Mooney, and M. Sorbero. 2000a. Physician Referral Rates: Style Without Much Substance? *Medical Care* 38:836–846.

Franks, P., G. C. Williams, J. Zwanziger, et al. 2000b. Why do Physicians Vary so Widely in Their Referral Rates? *Journal of General Internal Medicine* 15:163–168.

Franks, P., J. Zwanziger, C. Mooney, et al.1999. Variations in Primary Care Physician Referral Rates. *Health Services Research* 34:323–329.

Gazelle, G. S., E. F. Halpern, H. S. Ryan, et al. 2007. Utilization of Diagnostic Medical Imaging: Comparison of Radiologist Referral Versus Same-Specialty Referral. *Radiology* 245:517–522.

Goel V., K. Iron, and J. I. Williams. 1997. Enthusiasm or Uncertainty: Small Area Variations in the use of Mammography Services in Ontario, Canada. *Journal of Epidemiology & Community Health* 51:378–382.

Goldstein J. N., C. A. Camargo, Jr., A. J. Pelletier, et al. 2006. Headache in United States Emergency Departments: Demographics, Work-up and Frequency of Pathological Diagnoses. *Cephalalgia* 26:684–690.

Government Accountability Office (U.S.GAO). 2008. Medicare Part B Imaging Services: Rapid Spending Growth and Shift to Physician Offices Indicate Need for CMS to Consider Additional Management Practices GAO-08-462.

Greene, B. R., J. Barlow, and C. Newman. 1996. Ambulatory Care Groups and the Profiling of Primary Care Physician Resource Use: Examining the Application of Case Mix Adjustments. *Journal of Ambulatory Care Management* 19:86–89.

Grytten, J., and R. Sorensen. 2003. Practice Variation and Physician-Specific Effects. *Journal of Health Economics* 22:403–418.

Hartley, R. M., A. M. Epstein, C. M. Harris, et al. 1987. Differences in Ambulatory Test Ordering in England and America. Role Of Doctors' Beliefs and Attitudes. *The American Journal of Medicine* 82:513–517.

Health Services Research Group. 1992. Small-Area Variations: What are They and What do They Mean? Health Services Research Group. *Canadian Medical Association Journal* 146:467–470.

Hermer, L. D., and H. Brody 2010. Defensive Medicine, Cost Containment, and Reform. *Journal of General Internal Medicine*

Hillman, B. J. 2004. Self-Referral: The Time for Whining is Over. *Journal of the American College of Radiology* 1:795–796.

Hillman, B. J., C. A. Joseph, M. R. Mabry, et al. 1990. Frequency and Costs of Diagnostic Imaging in Office Practice--A Comparison of Self-Referring and Radiologist-Referring Physicians. *The New England Journal of Medicine* 323:1604–1608.

Hillman, B. J., G. T. Olson, P. E. Griffith, et al. 1992. Physicians' Utilization and Charges for Outpatient Diagnostic Imaging in a Medicare Population. *Journal of the American Medical Association* 268:2050–2054.

Hoey, J., J. M. Eisenberg, W. O. Spitzer, et al. 1982. Physician Sensitivity to the Price of Diagnostic Tests: A U.S.-Canadian analysis. *Medical Care* 20:302–307.

Holtgrave, D. R., F. Lawler, and S. J. Spann. 1991. Physicians' Risk Attitudes, Laboratory Usage, and Referral Decisions: The Case of an Academic Family Practice Center. *Medical Decision Making* 11:125–130.

Houchens, R., B. Chu, and C. Steiner. 2007. Hierarchical Modeling Using HCUP Data., 2007–01 edn.

Isaacs, D. M., J. Marinac, and C. Sun. 2004. Radiograph Use in Low Back Pain: A United States Emergency Department Database Analysis. *The Journal of Emergency Medicine* 26:37–45.

Katz, S. J., L. F. McMahon, and W. G. Manning. 1996. Comparing the Use of Diagnostic Tests in Canadian and US Hospitals. *Medical Care* 34:117–125.

Kessler, D. P., and M. B. McClellan. 2002. How Liability Law Affects Medical Productivity. *Journal of Health Economics* 21:931–955.

Kessler, D. P., N. Summerton, and J. R. Graham. 2006. Effects of the Medical Liability System in Australia, the UK, and the USA. *Lancet* 368:240–246.

Knottnerus, J. A., J. Joosten, and J. Daams. 1990. Comparing the Quality of Referrals of General Practitioners with High and Average Referral Rates: An Independent Panel Review. *British Journal of General Practice* 40:178–181.

Kramer, A. A., and J. E. Zimmerman. 2007. Assessing the Calibration of Mortality Benchmarks in Critical Care: The Hosmer-Lemeshow Test Revisited. *Crit Care Med* 35:2052–2056.

Kravitz, R. L., and S. Greenfield. 1995. Variations in Resource Utilization Among Medical Specialties and Systems of Care. *Annual Review of Public Health* 16:431–445.

Landon, B. E., J. Reschovsky, M. Reed, et al. 2001. Personal, Organizational, and Market Level Influences on Physicians' Practice Patterns: Results of a National Survey of Primary Care Physicians. *Medical Care* 39:889–905.

Lasko, T. A., S. J. Atlas, M. J. Barry, et al. 2006. Automated Identification of a Physician's Primary Patients. *Journal of the American Medical Informatics Association* 13:74–79.

Lawthers, A. G., A. R. Localio, N. M. Laird, et al. 1992. Physicians' Perceptions of the Risk of Being Sued. *Journal of Health Politics, Policy and Law* 17:463–482.

Leape, L. L., R. E. Park, D. H. Solomon, et al. 1990. Does Inappropriate Use Explain Small-Area Variations in the Use of Health Care Services? *Journal of the American Medical Association* 263:669–672.

Levy, G., A. Blachar, L. Goldstein, et al. 2006. Nonradiologist Utilization of American College of Radiology Appropriateness Criteria in a Preauthorization Center for MRI Requests: Applicability and Effects. *American Journal of Roentgenology* 187:855–858.

Long, M. J., K. M. Cummings, and K. B. Frisof. 1983. The Role of Perceived Price in Physicians' Demand for Diagnostic Tests. *Medical Care* 21:243–250.

Luke, D. A. 2004. *Multilevel Modeling*, 07-143 edn. Sage Publications, Thousand Oaks, Calif.

Lysdahl, K. B., and I. Borretzen. 2007. Geographical Variation in Radiological Services: A Nationwide Survey. *BMC.Health Services Research* 7:21.

Majeed, A., A. B. Bindman, and J. P. Weiner. 2001a. Use of Risk Adjustment in Setting Budgets and Measuring Performance in Primary Care I: How it Works. *British Medical Journal* 323:604–607.

Majeed, A., A. B. Bindman, and J. P. Weiner. 2001b. Use of Risk Adjustment in Setting Budgets and Measuring Performance in Primary Care II: Advantages, Disadvantages, and Practicalities. *British Medical Journal* 323:607–610.

Medicare Payment Advisory Commission. 2003. MedPAC: Variation and Inovation in Medicine. *Report to congress, June.*

Medicare Payment Advisory Commission. 2005. MedPAC Recommendations on Imaging Services. *Testimony Before Congress, March 17*.

Medicare Payment Advisory Commission. 2006. MedPAC Recommendations on Imaging Services. *Testimony Before Congress, July 18*.

Miller, W. L., R. R. McDaniel, Jr., B. F. Crabtree, et al. 2001. Practice Jazz: Understanding Variation in Family Practices Using Complexity Science. *The Journal of Family Practice* 50:872–878.

Mitchell, J. M. 2007. The Prevalence of Physician Self-Referral Arrangements After Stark II: Evidence from Advanced Diagnostic Imaging. *Health Affairs (Millwood)* 26:w415–w424.

Mold, J. W., and H. F. Stein. 1986. The Cascade Effect in the Clinical Care of Patients. *The New England Journal of Medicine* 314:512–514.

Mort, E. A., J. N. Edwards, D. W. Emmons, et al. 1996. Physician Response to Patient Insurance Status in Ambulatory Care Clinical Decision-Making. Implications for Quality of Care. *Medical Care* 34:783–797.

National Council on Radiation Protection and Measurements. 2009. NCRP Report No. 160: Ionizing Radiation Exposure of the Population of the United States. *Journal of Radiological Protection* 29:465.

Neilson, E. G., K. B. Johnson, S. T. Rosenbloom, et al. 2004. The Impact of Peer Management on Test-Ordering Behavior. *Annals of Internal Medicine* 141:196–204.

Nyweide, D. J., W. B. Weeks, D. J. Gottlieb, et al. 2009. Relationship of Primary Care Physicians' Patient Caseload with Measurement of Quality and Cost Performance. *The Journal of the American Medical Association* 302:2444–2450.

O'Neill, L., and J. Kuder. 2005. Explaining Variation in Physician Practice Patterns and Their Propensities to Recommend Services. *Medical Care Research and Review* 62:339–357.

Otero, H.J., S. Ondategui-Parra, E. M. Nathanson, et al. 2006. Utilization Management in Radiology: Basic Concepts and Applications. *Journal of the American College of Radiology* 3:351–357.

Parchman, M. L. 1995. Small Area Variation Analysis: A Tool for Primary Care Research. *Family Medicine Journal* 27:272–276.

Pekoz, E. A., M. Shwartz, L. I. Lezzoni, et al. 2003. Comparing the Importance of Disease Rate Versus Practice Style Variations in Explaining Differences in Small Area Hospitalization Rates for Two Respiratory Conditions. *Statistics in Medicine* 22:1775–1786.

Pezzin, L. E., P. M. Keyl, and G. B. Green. 2007. Disparities in the Emergency Department Evaluation of Chest Pain Patients. *Academic Emergency Medicine* 14:149–156.

Pham, H. H., G. C. Alexander, and A. S. O'Malley. 2007. Physician Consideration of Patients' Out-of-Pocket Costs in Making Common Clinical Decisions. *Archives of Internal Medicine* 167:663–668.

Phelps, C. E., C. Mooney, A.I. Mushlin, et al. 1994. Doctors Have Styles-- and They Matter! *Working paper.*

Phillips, K. A., K. R. Morrison, R. Andersen, et al. 1998. Understanding the Context of Healthcare Utilization: Assessing Environmental and Provider-Related Variables in the Behavioral Model of Utilization. *Health Services Research* 33:571–596.

Quintana, J. M., D. Goldmann, and C. Homer. 1997. Social Disparities in the use of Diagnostic Tests for Children with Gastroenteritis. *International Journal for Quality in Health Care* 9:419–425.

Raudenbush, S. W., and A. S. Bryk. 2002. *Hierarchical Linear Models Applications And Data Analysis Methods*, 2nd ed edn. Sage Publications, Thousand Oaks.

Reinhardt, U. E. 1999. The Economist's Model of Physician Behavior. *Journal of the American Medical Association* 281:462–465.

Restuccia, J., M. Shwartz, A. Ash, et al. 1996. High Hospital Admission Rates and Inappropriate Care. *Health Affairs (Millwood)* 15:156–163.

Rosen, M. P., R. B. Davis, and L. G. Lesky 1997. Utilization of Outpatient Diagnostic Imaging. Does the Physician's Gender Play A Role? *Journal of General Internal Medicine* 12:407–411.

Rosenthal, D. I., J. B. Weilburg, T. Schultz, et al. 2006. Radiology Order Entry with Decision Support: Initial Clinical Experience. *Journal of the American College of Radiology* 3:799–806.

Salem-Schatz, S., G. Moore, M. Rucker, et al. 1994. The Case for Case-Mix Adjustment in Practice Profiling. When Good Apples Look Bad. *Journal of the American Medical Association* 272:871–874.

Schulman, K. A., J. A. Berlin, W. Harless, et al. 1999. The Effect of Race and Sex on Physicians' Recommendations for Cardiac Catheterization. *The New England Journal of Medicine* 340:618–626.

Shen, J., R. Andersen, R. Brook, et al. 2004. The Effects of Payment Method on Clinical Decision-Making: Physician Responses To Clinical Scenarios. *Medical Care* 42:297–302.

Sirovich, B., P. M. Gallagher, D. E. Wennberg, et al. 2008. Discretionary Decision Making by Primary Care Physicians and the Cost of U.S. Health Care. *Health Affairs (Millwood)* 27:813–823.

Sirovich, B. E., D. J. Gottlieb, H. G. Welch, et al. 2006. Regional Variations in Health Care Intensity and Physician Perceptions of Quality of Care. *Annals of Internal Medicine* 144:641–649.

Sistrom, C.L. 2009. The Appropriateness of Imaging: A Comprehensive Conceptual Framework. *Radiology* 251:637–649.

Sistrom, C.L., P. A. Dang, J. B. Weilburg, et al. 2009. Effect of Computerized Order Entry with Integrated Decision Support on the Growth of Outpatient Procedure Volumes: Seven-Year Time Series Analysis. *Radiology* 251:147–155.

Sloan, F.A., and J. J. Shadle. 2009. Is there Empirical Evidence for "Defensive Medicine"? A reassessment. *Journal of Health Economics* 28:481–491.

Smith-Bindman, R. , J. Lipson, R. Marcus, et al. 2009. Radiation Dose Associated with Common Computed Tomography Examinations and the Associated Lifetime Attributable Risk of Cancer. *Archives of Internal Medicine* 169:2078–2086.

Smulowitz, P. B., L. Ngo, and S. K. Epstein. 2009. The Effect of a CT and MRI Preauthorization Program on ED Utilization. *The American Journal of Emergency Medicine* 27:328–332.

Solomon, D. H., H. Hashimoto, L. Daltroy, et al. 1998. Techniques to Improve Physicians' Use of Diagnostic Tests: A New Conceptual Framework. *Journal of the American Medical Association* 280:2020–2027.

Sood, R., A. Sood, and A. K. Ghosh. 2007. Non-evidence-based Variables Affecting Physicians' Test-ordering Tendencies: A Systematic Review. *The Netherlands Journal of Medicine* 65:167–177.

Stahl, F. W. 1969. *The Mechanics of Inheritance*, 2d edn. Prentice-Hall, Englewood Cliffs, N.J.

Stano, M. 1991. Further Issues in Small Area Variations Analysis. *Journal of Health Politics, Policy and Law* 16:573–588.

Stano, M. 1993. Evaluating the Policy Role of the Small Area Variations and Physician Practice Style Hypotheses. *Health Policy* 24:9–17.

Thrall, J. H. 2005a. Reinventing Radiology in the Digital Age. Part II. New Directions and New Stakeholder Value. *Radiology* 237:15–18.

Thrall, J. H. 2005b. Reinventing Radiology in the Digital Age. Part III. Facilities, Work Processes, and Job Responsibilities. *Radiology* 237:790–793.

Thrall, J. H. 2005. Reinventing Radiology in the Digital Age. Part I. The All-digital Department. *Radiology* 236:382–385.

Tucker, A. M., J. P. Weiner, S. Honigfeld, et al. 1996. Profiling Primary Care Physician Resource Use: Examining the Application of Case Mix Adjustment. *Journal of Ambulatory Care Management* 19:60–80.

Verrilli, D., and H. G. Welch. 1996. The Impact of Diagnostic Testing on Therapeutic Interventions. *Journal of the American Medical Association* 275:1189–1191.

Wasiak, R., G. S. Pransky, and S. J. Atlas. 2008. Who's in Charge? Challenges in Evaluating Quality of Primary Care Treatment for Low Back Pain. *Journal of Evaluation in Clinical Practice* 14:961–968.

Welch, H. G., M.E. Miller, and W. P. Welch. 1994. Physician Profiling. An Analysis of Inpatient Practice Patterns in Florida and Oregon. *The New England Journal of Medicine* 330:607–612.

Welch, W. P., M.E. Miller, H.G. Welch, et al. 1993. Geographic Variation in Expenditures for Physicians' Services in the United States. *The New England Journal of Medicine* 328:621–627.

Wennberg, D. E., J. D. Dickens, Jr., L. Biener, et al. 1997. Do Physicians Do What They Say? The Inclination to Test and its Association with Coronary Angiography Rates. *Journal of General Internal Medicine* 12:172–176.

Wennberg, J., and A. Gittelsohn. 1973. Small Area Variations in Health Care Delivery. *Science* 182:1102–1108.

Wennberg, J. E. 1987a. Population Illness Rates Do Not Explain Population Hospitalization Rates. A Comment on Mark Blumberg's Thesis that Morbidity Adjusters are Needed to Interpret Small Area Variations. *Medical Care* 25:354–359.

Wennberg, J. E. 1987b. The Paradox of Appropriate Care. *Journal of the American Medical Association* 258:2568–2569.

Wennberg, J. E. 1993. Future Directions for Small Area Variations. *Medical Care* 31:YS75–YS80.

Wennberg, J. E. 2002. Unwarranted Variations in Healthcare Delivery: Implications for Academic Medical Centres. *British Medical Journal* 325:961–964.

Wennberg, J. E. 2004a. Practice Variation: Implications for Our Health Care System. *Managed Care* 13:3–7.

Wennberg, J. E. 2004b. Practice Variations and Health Care Reform: Connecting the Dots. *Health Affairs (Millwood)* Suppl Web Exclusives, VAR140–VAR144.

Wennberg, J. E., B. A. Barnes, and M. Zubkoff. 1982. Professional Uncertainty and the Problem of Supplier-Induced Demand. *Social Science & Medicine* 16:811–824.

Westert, G. P., and P. P. Groenewegen. 1999. Medical Practice Variations: Changing the Theoretical Approach. *Scandinavian Journal of Public Health* 27:173–180.

Whiting, P., M. Toerien, I. de Salis, et al. 2007. A Review Identifies and Classifies Reasons for Ordering Diagnostic Tests. *Journal of Clinical Epidemiology* 60:981–989.

Williams, S. V., J. M. Eisenberg, L. A. Pascale, et al. 1982. Physicians' Perceptions about Unnecessary Diagnostic Testing. *Inquiry* 19:363–370.

Wu, G. 1996. The Strengths and Limitations of Expected Utility Theory. *Medical Decision Making* 16:9–10.

BIOGRAPHICAL SKETCH

After serving in the U.S. Army Signal Corps as a cryptographer from 1973-1977, Chris Sistrom, MD, MPH obtained his undergraduate degree in Computer Science (1980) from the University of Oregon in Eugene. He attended at Oregon Health Sciences University (MD in 1984) and completed radiology residency at the University of Virginia in 1988. He is now Associate Chairman of Radiology, Chief Information Officer for Radiology, and Associate Professor at the University of Florida, College of Medicine. Dr. Sistrom obtained an MPH degree in epidemiology and health policy in 2003 from the University of Florida, and is in final stages of a PhD in Health Services Research there. The topic of his dissertation is Imaging Utilization in Primary Care. The research goal is to quantify and model various factors that affect the intensity and mixture of outpatient imaging performed on primary care patients. The resulting models will be useful in practitioner profiling at institutional and regional levels. The eventual goal is to produce a ‘Map of Imaging’ along the lines of the Dartmouth Atlas of Healthcare and to create risk adjusted population based estimates of optimal and appropriate imaging utilization.