Theory and Measurement in the Study of Medical Practice Variation
Theory and Measurement in the Study of Medical Practice Variation

By: MATHEW MERCURI, B.Kin., M.Sc.

A Thesis
Submitted to the School of Graduate Studies
In Partial Fulfillment of the Requirements for the Degree
Doctorate of Philosophy

McMaster University © Copyright by Mathew Mercuri, April 2013
TITLE: Theory and Measurement in the Study of Medical Practice Variation

AUTHOR: Mathew Mercuri, B.Kin. (McMaster University), M.Sc. (McMaster University)

SUPERVISOR: Dr. Amiram Gafni

NUMBER OF PAGES: xiii, 154
Abstract

Variations in the rate of use of health care services (a.k.a. medical practice variations) have been described in the literature for over eighty years. The literature suggests three general sources of variation: patients, physicians, and environment. The relative influence of these sources and the specific mechanisms that produce observed variations are not well understood. This thesis presents four studies that identify and examine methodological issues that preclude our ability to understand the variation phenomenon.

It is commonly believed that the physician is in part responsible for observed variation in health care services use. However, determining the influence of the physician in this regard is problematic, as it is difficult to isolate the effect of the physician from that of the patient and environment (including available resources). The first study presented in this thesis suggests there is meaningful variation in treatment recommendations between physicians working in a common environment, even after controlling for important patient clinical characteristics. Next, I present an experiment that suggests that factors related to the patient’s unique social context might influence how the physician intends to manage a patient’s care. As variations studies do not measure or adjust for social context, this might indicate an important methodological limitation of those studies if indeed context is an important (and justifiable) determinant of what care the patient will receive.
Not all variation is necessarily bad. The third study I present explores how previous researchers discriminate between warranted and unwarranted variation. This study indicates that few researchers explicitly do so, and that a clear, consistent, and functional definition of unwarranted variation is lacking – a feature that potentially limits the interpretation of study results. The final study argues that traditional methods for examining regional variations are inadequate for informing health care managers because they examine variation in health care service use rather than needs.
Acknowledgments

This thesis would not have been possible without the help, support, and guidance of a number of colleagues and friends.

First and foremost I would like to thank my supervisor, Dr. Amiram Gafni, and my thesis committee members, Dr. Geoff Norman and Dr. Scott Watter. Their guidance demonstrates the importance of mentoring in graduate education. Much was learned regarding the intangible aspects of good academic inquiry, and how to navigate a career in an academic field—all by virtue of time spent in conversation with the aforementioned gentlemen over the past 5 years. I am grateful for the opportunity to have learned from each of them, and for their time they gave me over the course of my studies.

I would also like to thank those who helped with the studies that make up this thesis. I would like to thank Dr. Madhu Natarajan and Dr. James Velianou for their help and guidance in all things cardiology, providing me with access to the interventional cardiology registry, and for giving me insight into how physicians work and make decisions. I would like to thank Dr. Jonathan Sherbino for developing the emergency medicine clinical scenarios needed for our examination of the influence of social context on management decisions, and for providing me with a large group of eager participants for that study. My appreciation also goes to Sharyn Kreuger for her help in developing the electronic survey for that study, and for managing the collection of data. Thanks to Dr. Stephen Birch for his advice on both health care needs and organization of health care services.

I would like to express my gratitude to members of the Health Research Methodology Program and Global Health offices for their administrative help, counseling, and the interesting teaching opportunities they provided. Specific thanks go to Dr. Steven Hanna, Lorraine Carroll, Kristine Leadbetter, Kristina Vukelic, Dr. Andrea Baumann, and Laurie Kennedy.

A very special thanks goes to my friends at PERD. Somehow, despite your best efforts, I ended up as a part of your group (socially—maybe now we can work on it academically). Thank you to (soon to be Dr.) Mahan Kulasegaram, Dr. Lawrence Grierson, Dr. Kelly Dore, and Katie Zazulak.

I would like to thank my family and friends for their ongoing support. My special thanks goes to Dr. Catherine Kreatsoulas for her moral support in this endeavour.

Finally, I would like to thank the most important person of all—my wife, Diana Weir. Thank you for living with the uncertainty of where we might be when this was all done, and for being so patient.
Table of Contents

Title Page...........................................................................................................................................ii
Descriptive Note.......................................................................................................................................iii
Abstract....................................................................................................................................................iv
Acknowledgements...............................................................................................................................vi
Table of Contents..................................................................................................................................vii
List of Tables...........................................................................................................................................x
List of Figures..........................................................................................................................................x
List of Abbreviations..............................................................................................................................xi
Declaration of Academic Achievement.................................................................................................xii

Chapter 1: Introduction.................................................................................................................................1

Chapter 2: The Evolution of Medical Practice Variations Research: A Case of Insufficient Tools or the Wrong Blueprint?.................................................................8

2.1 A Tonsillectomy for some and not for others: Practice Variation Literature before 1940................................................................................................................................................10
2.2 Observed Variation (1940s-1970s), and Models of Health Service Use.....................................11
2.3 Wennberg’s “Small Area Variations”...............................................................................................16
2.4 “Uncertainty” in the Medical Evidence: The Underlying Reason for Variation?...................22
2.5 “Practice Style”: The Physician as the Source of Variation..........................................................25
2.6 Contextual/Environmental Sources of Variation........................................................................30
2.7 Examining Variation – The General Picture and Questions Unanswered..............................33
2.8 Preliminary Conclusions....................................................................................................................37

Chapter 3: An Even Smaller Area Variation: Differing Practice Patterns Among Interventional Cardiologists Within A Single High Volume Tertiary Cardiac Centre.................................................................40

3.1 Abstract.............................................................................................................................................42
3.2 Background.......................................................................................................................................42
3.3 Methods...........................................................................................................................................45
3.3.1 Study Site.....................................................................................................................................45
3.3.2 Population..................................................................................................................................46
3.3.3 Data Collection and Analyses.....................................................................................................47
3.4 Results...............................................................................................................................................50
3.5 Discussion..........................................................................................................................................52
3.5.1 Limitations..................................................................................................................................56
3.6 References........................................................................................................................................59
Chapter 4: The Influence of Patient Social Context in Determining
Emergency Medicine Physicians’ Management Decisions Relative
to Clinical Practice Guideline.........................................................64

4.1 Introduction..............................................................................64
4.2 Methods..................................................................................67
4.2.1 Participants...........................................................................67
4.2.2 Survey Instrument.................................................................68
4.2.3 Data Acquisition.................................................................69
4.2.4 Analysis...............................................................................70
4.3 Results......................................................................................70
4.4 Discussion...............................................................................72
4.4.1 Limitations..........................................................................78
4.5 Conclusion..............................................................................79

Chapter 5: Medical Practice Variations: What the Literature Tells Us (or Does
Not) About What Are Warranted and Unwarranted Variation........82

5.1 Abstract....................................................................................84
5.2 Introduction..............................................................................85
5.3 Literature Search......................................................................86
5.4 Sources of Practice Variation..................................................87
5.4.1 Types of Practice Variations................................................87
5.4.2 Uncovering the Sources of Practice Variations: Early Research and the
Role of the Physician.................................................................88
5.4.3 Criticism of the “Practice Style” Explanation and the Role of
Contextual Constraints..............................................................90
5.5 Are All Variations Bad?............................................................91
5.5.1 Suggested Frameworks of Unwarranted Variation in the Literature...92
5.5.2 Articles Implying a Framework............................................95
5.6 Discussion...............................................................................96
5.7 Conclusion..............................................................................98
5.8 References..............................................................................99

Chapter 6: Using Small Area Variations to Inform Healthcare Service
Planning: What Do We “Need” to Know?........................................108

6.1 Abstract....................................................................................110
6.2 Introduction..............................................................................111
6.3 Conceptual Framework..........................................................113
6.3.1 Demand, Need, Use, and Supply........................................113
6.3.2 Healthcare Policy Goals......................................................116
6.4 Small Area Variations: What they measure, and how they might (mis)inform Health care Policy
6.5 Modifying Small Area Methods to Inform Policy
6.6 Conclusions
6.7 References

Chapter 7: Discussion and Conclusions

7.1 Summary of Important Findings
7.2 Accounting for Health Care Needs
7.3 Accounting for Patient Context
7.4 The Physician as a Source of Variation
7.5 Overall Picture
7.6 Conclusions

References
List of Tables

Table 3-1: Clinical characteristics/demographics of patients in each of the three recommended treatment groups.................................................................63

Table 3-2: Adjusted odds of prescribing revascularization over medical therapy, and PCI over CABG for each physician in the study compared to that for the reference physician.......................................................................................63

Table 4-1: A brief summary of the survey cases and their associated social variables........................................................................................................80

Table 4-2: Participant demographics and characteristics of work environment. Rate of adherence to CPGs...........................................................................80

List of Figures

Figure 4-1: Rate of adherence to CPGs for both experienced and resident physicians, in both the SV and no SV conditions.........................................................81

Figure 6-1: Use (U) is the intersection of the Supply (S) and Demand (D) schedules.......................................................................................................................128
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CABG</td>
<td>Coronary Artery Bypass Graft Surgery</td>
</tr>
<tr>
<td>CAD</td>
<td>Coronary Artery Disease</td>
</tr>
<tr>
<td>COPD</td>
<td>Chronic Obstructive Pulmonary Disease</td>
</tr>
<tr>
<td>CPG</td>
<td>Clinical Practice Guideline</td>
</tr>
<tr>
<td>CT</td>
<td>Computed Tomography</td>
</tr>
<tr>
<td>EBM</td>
<td>Evidence Based Medicine</td>
</tr>
<tr>
<td>ED</td>
<td>Emergency Department</td>
</tr>
<tr>
<td>EM</td>
<td>Emergency Medicine</td>
</tr>
<tr>
<td>HAART</td>
<td>Highly Active Antiretroviral Therapy</td>
</tr>
<tr>
<td>LAD</td>
<td>Left Anterior Descending Artery</td>
</tr>
<tr>
<td>LVEF</td>
<td>Left Ventricular Ejection Fraction</td>
</tr>
<tr>
<td>MI</td>
<td>Myocardial Infarction</td>
</tr>
<tr>
<td>MOHLTC</td>
<td>Ministry of Health and Long Term Care</td>
</tr>
<tr>
<td>NICE</td>
<td>National Institute for Health and Clinical Excellence</td>
</tr>
<tr>
<td>PCI</td>
<td>Percutaneous Coronary Intervention</td>
</tr>
<tr>
<td>PVD</td>
<td>Peripheral Vascular Disease</td>
</tr>
<tr>
<td>RAWP</td>
<td>Resource Allocation Working Party</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomized Controlled Trial</td>
</tr>
<tr>
<td>SAV</td>
<td>Small Area Variation</td>
</tr>
<tr>
<td>SV</td>
<td>Social Variable</td>
</tr>
<tr>
<td>VD</td>
<td>Vessel Disease</td>
</tr>
</tbody>
</table>
Declaration of Academic Achievement

I am the principal author and sole writer of the contents of this thesis. Where the thesis contains multi-authored publications, my specific role is described, below. All published work has been reproduced for this thesis with permission.

Chapter 3: An Even Smaller Area Variation: Differing Practice Patterns Among Interventional Cardiologists Within A Single High Volume Tertiary Cardiac Care Centre.


Candidate’s role: Conceived of the study and design, performed statistical analyses. Drafted the manuscript and was primary author of the publication.

Chapter 4: The Influence of Patient Social Context in Determining Emergency Medicine Physicians’ Management Decisions Relative to Clinical Practice Guidelines

Unpublished Pilot Study

Candidate’s role: Co-designer of the study, performed statistical analyses. Drafted the manuscript.

Chapter 5: What the Literature Tells Us (or Does Not) About What Are Warranted and Unwarranted Variations


Candidate’s role: Devised and conducted a systematic literature review. Drafted the manuscript and was primary author of the publication.
Chapter 6: Using Small Area Variations to Inform Healthcare Service Planning: What Do We “Need” to Know?


Candidate’s role: Performed a literature review and synthesis. Drafted the manuscript and was the primary author of the publication.
Chapter 1: Introduction

Observed variations in the rate of use of healthcare services or resources are often reported in the healthcare literature (e.g. Glover, 1938; Wennberg & Gittelsohn, 1973; Paul-Shaheen, Clark, and Williams, 1987; Jaffer et al., 2010; O'Connor et al., 1999; Ugnat and Naylor 1994). These variations can be between seemingly similar groups of patients linked to different individual physicians (i.e. practice variations), groups of physicians, or hospitals/geographic areas treating a seemingly similar mix of patients (i.e. “small area variations”). Some might argue that, as these variations cannot be fully explained by differences in patient health status, observed variations raise questions about the quality of care, efficiency of limited resource use, and/or equity in the provision of services (e.g. Cohen et al., 1992; Wennberg, 2011). Therefore, healthcare researchers and managers have sought to both better understand the causes of these variations, and minimize their (i.e. variation) occurrence.

Over the past few decades we have seen a rapid increase in the number of research studies reporting variation in the rate of use of healthcare services. As the existence of variations was becoming apparent in the literature, researchers surmised that closing the knowledge gap regarding the best course of treatment for a given medical condition, for example through the use of clinical trials and comparative effectiveness research, might result in a drastic reduction in the magnitude of variation between individual or groups of physicians (Heasman and
Carstairs, 1971; Wennberg, 1984, Wennberg, 2011). Shortly after variations studies became commonplace in the mainstream healthcare literature (Paul-Shaheen et al., 1987), clinical scientists began establishing formal techniques for determining the comparative effectiveness of available therapies, and thus the best course of treatment given a specific medical condition, through the paradigm of “Evidence-Based Medicine” (EBM) (The Evidence-Based Medicine Working Group, 1992). At that time, both influential practice variations researchers, such as Wennberg (1984), and advocates for EBM shared the belief that many management decisions in healthcare were inappropriately determined by characteristics of the individual physician, rather than dictated by the medical evidence1. Based on the timing of events and the common desire to elevate the influence of medical evidence on management decisions, one could argue that observed practice variations, especially those attributing variation to physician level factors, were the impetus for EBM, itself a highly influential and important philosophy in medical practice, research, and health service planning.

However, the effectiveness of strategies to minimize aforementioned variation is predicated on our ability to correctly identify their existence, and determine their causes (i.e. the sources of variation). For example, if evidence demonstrates that variation is a result of a knowledge deficit or professional uncertainty about the best course of treatment, then strategies to reduce variation

---

1 In some cases patient preference might also dictate the best course of action. For example, preference might take precedence when there is equipoise among the available therapies (see Wennberg, 2004; Guyatt, Rennie, Meade, & Cook, 2008).
must target those factors (e.g. clinical practice guidelines (CPGs)). If it were the case that variations were caused by factors related to the patient’s context, despite clear medical evidence regarding the benefit of one therapy over another, any strategy that targets “knowledge (translation) gaps” would be ineffective.

It seems that a common idea in the healthcare literature is that if variation happens it is because physicians or patients do not know what is the right course of action for a given medical condition. It might be that the physician or patient is not aware of what the medical evidence reports to be the best therapy, or that such evidence is weak or does not exist. However, despite two decades of EBM (including clinical trials, CPGs, and knowledge translation strategies), variations in the rate of use of healthcare services continue to be reported. While it is possible that EBM has done much to reduce variation, it could be that such variation is only in part due to insufficient knowledge regarding the best available therapy. It is reasonable to believe that variation might be caused by characteristics of the patient, the physician, and the environment in which care is provided (including available resources), independent of what is known to be the “evidence-based” best course of action. On the other hand, it may be that how researchers examine and define variation is problematic in our understanding of this phenomenon. It could be that what we consider to be variation when using current methods in fact reflects appropriate differences in utilization based on good use of resources to attain desired outcomes (e.g. maximize health outcomes to the community, efficient use of resources, etc.).
Variations studies yield information on differences in the rate of use between two or more defined groups of patients thought to be similar in all respects related to their health status and access to care. However, variation is simply a statistic (Tanenbaum, 2013). The meaning of this statistic depends on the method by which it is derived. The purpose of this thesis is to examine the methodological nature of variations studies in healthcare services.

I begin with a brief history of healthcare variations research in chapter 2. There I describe major studies examining variations in the rate of use of healthcare services, starting with Glover’s work on differing tonsillectomy rates among British school children in the 1930s, through Wennberg and colleagues work on “small area variations” starting in 1973, ending with more current studies. Throughout I will review proposed frameworks to examine healthcare service use and variations in such. This chapter will also include a review of the evidence for and against patients, physicians, and the context/environment as general sources of variation. I will conclude chapter 2 by proposing three methodological issues revealed from the review of literature on variations that require a more extensive examination.

Any discussion on variation in healthcare service use begins with the decision maker, namely the physician managing care. It is a somewhat widespread belief among variations researchers, especially in the early years of the field, that it is the physicians themselves who are most responsible for why seemingly similar groups of patients receive one form of therapy or another at different rates. The extent to which this is true is debated in the literature (as will be demonstrated in
Part of this debate is related to how researchers designed and executed those studies that attribute variations to physicians. Chapter 3 presents a research study that examines if meaningful variations in the rate of use of three healthcare services exist between physicians when patient and context/environment related factors are constant. Specifically, this study examines between physician differences in the rate of use of angioplasty, cardiac surgery, and medical therapy among patients with diagnosed coronary artery disease treated at a single, high volume cardiac centre, within a publicly funded healthcare system. This study provides an added twist to traditional methods in that it examines what the physicians recommended should be the course of treatment rather than what the patient ultimately received.

Professional uncertainty regarding the best course of care is seen as a culprit in generating variation. Clinical practice guidelines were developed to assist physicians in applying the best evidence-based care for a given medical condition, thus reducing professional uncertainty. Thus, where evidence-based CPGs exist it seems unreasonable that patients with similar health status would be differently managed. However, patients bring more than just their health status when presenting to physicians for care. There is no guarantee that, given his or her situation, the patient will tolerate what the evidence suggests is the best course of care. Features unique to the individual patient’s context have not been considered in previous studies examining practice variation. First one must determine if these features influence the physicians’ recommendations for care. Chapter 4 examines
the role of the patient’s “social context” in determining the course of care he or she will receive. Specifically, this is done in the context of management decisions among emergency medicine specialist physicians for cases where CPGs exist. As a subsequent focus, this study will examine if experience is a factor in how physicians respond to patient social context.

Variation in the use of healthcare services is not necessarily bad. As it is doubtful one would wish to eliminate variations that result in overall better care or use of resources, it is important that one identifies those variations that affect attainment of the desired outcomes. Thus, discriminating between “good” and “bad” variations should be part of the framework researchers employ when examining variation. Chapter 5 provides a systematic review of how researchers define “unwarranted” variations, and an examination of how such definitions relate to their methodology when identifying and explaining variations.

Chapter 6 focuses on the policy relevance of information derived from variations studies, specifically those using “small area” methods. Proponents of small area variations research argue that as observed variations speak to issues of efficiency and effectiveness of healthcare services, policy makers and healthcare managers should be made aware of their existence so that steps can be taken to minimize their magnitude and/or negative effects on overall population health outcomes. This chapter describes common healthcare policy goals, what small area methods measure, and examines whether these are aligned. A significant methodological issue is revealed. Specifically, healthcare policy is typically
concerned with allocation of resources to maximize health outcomes on the basis of needs, whereas small area methods measure variations in use.

I conclude this thesis with a general discussion regarding how these methodological issues together limit our ability to identify important variations, and determine their causes. This will include a discussion regarding what can be done to improve the way researchers engage the phenomenon - most notably the need for a functional theory, and what aspects should be included in that theory and any derived frameworks for investigation (i.e. what should be measured, and how it should be measured).
Chapter 2: The Evolution of Medical Practice Variations Research: A Case of Insufficient Tools or the Wrong Blueprint?

Consider the following scenario: similar groups of individuals (based on known characteristics) living in two different geographic areas are each diagnosed with two vessel coronary artery disease. However, in one region the rate of angioplasty among these individuals is twice that in the other region. What is described in this scenario is known as a medical practice variation (more specifically a geographic or “small area” variation). In a more general sense, a practice variation can be defined as the difference in the use of health services or resources between individual or groups of patients who (all) share a (seemingly) similar health status or medical condition.²

Practice variations have received a great deal of attention in the healthcare literature over the past few decades. Such variations cause alarm among many healthcare researchers, planners, and providers, as their existence raises concerns that to them: 1) some patients are not receiving the best care, and/or 2) resources are used inefficiently in some areas, indicating opportunities for potential cost savings in the overall provision of care where there is suspicion of overuse of a

---
² In some situations throughout this document I prefer “variation in health service/resource use” rather than “practice variation” as a general term for the phenomenon of interest – “practice variation” seemingly implies the physician as the culprit (whether this is truly the case is still to be determined as will become clear throughout this review).
given medical diagnostic, intervention, or other healthcare resource. Furthermore, while some variation might be expected in any system (i.e., some variations are “justified”, or are “random error”), healthcare or otherwise, it is the magnitude of the variation that causes concern among stakeholders in that it is unexpected given the perceived or measured level of difference in resource availability and/or health status between populations. Thus, healthcare researchers have taken great interest in both revealing the existence of these variations, and determining their causes.

Variations research has evolved over the past eight decades. However, what becomes apparent when reviewing the research to date is that studies examining variation have origins in simple observation, with posteriori adjustment for confounding factors and explanations of cause, rather than based on theory and grounded in an academic discipline – for the most part a characteristic of such studies even today. Likewise, a number of measurement issues become apparent when reviewing the literature. Both the lack of theory and proper measurement (perhaps due to a lack of theory to guide measurement) constitute important methodological issues that may limit our ability to understand the variation phenomenon. The purpose of this review is to trace the history and the evolution (or potentially lack thereof) of studies examining variations in the use of healthcare

---

3 This argument is made when outcomes, for example, mortality rates, appear to be similar between geographic areas despite differing levels of resource use. Reason would suggest that if providers can achieve the same results with fewer resources, than all populations should adopt that model of resource use. However, the extent to which this is true is debatable.
services, so as to determine the current state of knowledge in this area and elucidate gaps in the current methods.

2.1 A Tonsillectomy for some and not for others: Practice Variation Literature before 1940

Although research on practice variations gained some notoriety through Wennberg's work in the 1970s (Wennberg and Gittelsohn, 1973), their existence in the healthcare literature can be traced to at least four decades earlier. While examining the ever-increasing rate of tonsillectomy in British school children, Glover and Wilson (1932) revealed numerous categories of variation. For example, their review of the literature showed the tonsillectomy rate in Germany to be only a fraction of that among children in England and Wales. Substantial variation in the rate of tonsillectomy was also apparent between samples of children from various American States (i.e. California, Minnesota, and Missouri), which again was in some states only a fraction of that seen in British school children (Glover and Wilson, 1932). Glover and Wilson also found variation between males and females, and between public and private school children, implying that the use of tonsillectomy surgery may be associated with the ability to pay for service. These variations were unlikely to be explained by between population differences in health status, as the rate of use in most regions was substantially higher than the incidence of infection where tonsillectomy is indicated.
A subsequent and more focused examination of the regional rates of tonsillectomy among children in England and Wales would confirm the practice variation implied in Glover and Wilson's original report on tonsillectomy use (Glover, 1938). Although Glover would find “puzzling” the comparatively higher incidence of tonsillectomy among the “children of the well-to-do” despite no demonstrated class-level differences in the incidence of tonsillitis, he would ultimately conclude that a “study of the geographical distribution in elementary school children discloses no correlation between the rate of incidence and any impersonal factor, such as over-crowding, poverty, bad housing, or climate. Incidence is not correlated with the general efficiency of the school medical and dental services of the area. In fact it defies any explanation, save that of variations of medical opinion on the indications for operation” (Glover, 1938; p.1234). Glover’s conclusion, along with the implied “overuse”\(^4\) of medical intervention (i.e. tonsillectomy) in many regions, would set the tone for much of the later work in medical practice variations.

2.2 Observed Variation (1940s-1970s), and Models of Health Service Use

Practice variations would continue to be revealed in the next few decades after Glover’s observed variation in tonsillectomy rates (1938). For example,\(^4\)

\(^4\) Overuse of tonsillectomy is implied in two ways in Glover’s studies. First, the subsequent rate of infection/disease (including those nasopharyngeal, influenza, infectious disease, etc.) was no different between tonsillectomized and non-tonsillectomized children (Glover, 1932). Second, a drastic reduction in the rate of tonsillectomy in some regions did not produce “unsatisfactory results” where “tonsils showed definite enlargement” (Glover, 1938).
Ciocco, Altman, and Truan (1952) would observe variations between Pennsylvania counties in the annual per capita use of physician services. This was in part attributed to the per capita volume of physicians, although the authors did admit this to be a weak association. Around the same time, Dunlop, Inch, and Paul (1953) observed variation in the mean monthly prescription rate (per 1000 individuals) among Scottish executive council (i.e. geographic) areas. To explain this variation, the authors offer differential morbidity and/or patient attitude towards disease as potential factors. Furthermore, as the prescription rate was associated with the population density of the geographic area, access to a physician was also seen as an important factor, or as the authors eloquently state, “in urban areas the practitioner is usually conveniently available for a consultation and a prescription, often for very trivial maladies, in contrast to the Highlands and Islands of Scotland, where moor and fen, crag and torrent, and even stormy seas may intervene between the patient and doctor” (Dunlop et al. 1953; p.697). Likewise, Lembke (1952) surmised that accessibility to hospital care (e.g. rural vs. urban living) is a factor in explaining age- and sex-adjusted variations in appendectomy rates across counties in western New York State. In other words, the patients’ circumstances (including those non-medical) may play a role in determining the rate in which they seek care, which in aggregate may result in observed variation in the use of healthcare services (assuming systematic differences in these circumstances between populations).

The idea that differential use of healthcare services may be due to patient context/circumstances gained some traction in the 1960s and early 1970s. Studies
and models espousing features of the individual, the community, and the available resources, rather than medical opinion alone, in determining the use of healthcare services were prominent in the healthcare literature at this time. Andersen and Newmen (1973) proposed what is perhaps the best known among the conceptual models of healthcare service use. Andersen and Newmen posited that use of healthcare services is determined by 1) both societal norms in values or beliefs and the impact of technology (Societal Determinants), 2) the amount and type of available resources and their organization (Health Services System), and 3) characteristics of the individual's circumstance (Individual Determinants). Each component can influence the others. For example, societal norms in values can determine how healthcare resources will be distributed and how the provision of care will be financed. Individual determinants go beyond health status. They include the individual's 1) predisposition to using healthcare services related to demographics, social structure, and beliefs about health, 2) ability to secure services, based on income, insurance, availability of healthcare providers, and location of those services relative to his or her home, and 3) health status, both perceived and diagnosed (Andersen and Newmen 1973).

Ro (1969) also offered a multifaceted framework for determining healthcare service use, integrating both medical and contextual factors. Whereas the Andersen and Newmen model examined health service use from a sociologic perspective, Ro’s model was grounded in economic theories of consumption and consumer behaviour. In the traditional (behavioural economic) sense, consumer behaviour in
the purchase of goods or use of services is determined by “choice-conditioning” factors (i.e. personal or situational factors, including health status in the case of healthcare services). However, when considering the use of healthcare services, the consumer’s choice to use such services is also dependent on the judgment of a physician (or other healthcare professional) who must authorize its purchase. Thus, in the realm of healthcare, consumer behaviour in the use of services is dependent on the physician’s response to the consumer’s choice-conditioning factors. Ro contends that use of healthcare services is somewhat more complex in that behaviour is shaped not only by the interaction between the patient and the physician, but rather by the “composite effect of the joint interaction among physicians, patients and hospital” (Ro, 1969; p.296). In other words, the patient’s medical condition and personal circumstances/values shape his or her intentions concerning health service use. These intentions are then modified based on both the physician’s response to the patient’s circumstance and intentions, and characteristics of the (hospital) environment. Furthermore, how the physician

---

5 In the introduction this manuscript, Ro states that traditionally the “medical profession has promoted the idea that, given one’s medical condition, the type and amount of hospital services provided are dictated by the “technological imperatives”. Any variations in hospital use which are not explained by the differences in patients’ medical conditions are attributed to the differences in the medical philosophies of the attending physicians and to “extra-medical” factors such as age and sex of the patient” (Ro, 1969; p. 296). This is similar to the contention of Glover (1938), with the added feature of “extra-medical” factors. However, Ro would state that recent (to his 1969 paper) studies have emerged to question this tradition, whereby two theories have emerged; one that accounts for economic theories of behaviour and consumption, and another that denies these theories, placing the use of services as determined by technological imperatives – medical conditions of the potential users dictate the product mix of each hospital. Ro’s model integrates both theories.
responds to the patient in part is determined by the physician’s personal and professional background.

Contemporaneous with the models of healthcare service use mentioned above was empirical evidence to support the influence contextual and/or personal factors specific to the individual (group of) patient(s) can have on determining use. Many of these factors are the same or similar to those outlined by Ro (1969), and Andersen and Newmen (1973). For example, Suchman (1964) showed that differential participation in health programs was associated with the level of estrangement from mainstream middle-class American society, brought on by the form of the social group. Where individuals belonged to highly ethnocentric and socially cohesive groups, those individuals more likely displayed low knowledge about disease and skepticism towards professional medical care. Weiss and Greenlick (1970) demonstrated that both the geographic distance from healthcare providers and the social class of the individual needing care affect initiation of contact with a physician. However, while these studies show differential use on the basis of non-medical or contextual factors, the extent to which this translates to variation between physicians, hospitals or geographic regions is contingent on systematic variation in these factors on the basis of the defined populations, geographic or otherwise.

When examining observed variations in use at that time, similar support was found for the influence of contextual and non-medical factors, in addition to intrinsic physician factors (e.g. opinion). For example, structural mechanisms related to how
care is provided and available resources, such as bed availability and staffing ratios were thought to explain variations in the per capita admission rate, length of stay, and per diem cost between the United States and Canada (Andersen and Hull, 1969). Similar factors (e.g. bed availability, healthcare structure, number of physicians per capita) were shown to explain geographic variations (in rate of use, length of stay, etc.) in other jurisdictions\(^6\) as well (Lewis, 1969; Vayda 1973, Feldstein, 1964; Bunker, 1970). In considering the patient as a source of variation, socioeconomic (i.e. non-medical) factors, most notably the patient’s occupation, were shown to explain some of the variation in appendectomy rates between residential districts in Hanover Germany (Lichtner and Pflanz, 1971). Physician related factors in part accounted for observed variations in healthcare service use as well. Mechanisms for such variations included when and where the physicians trained, how they were taught, their specialty/board certification, how they were compensated, and their service delivery model (group vs. solo practice) (Bunker, 1970; Freeborn, Baer, Greenlick, & Bailey, 1972; Childs and Hunter, 1972; Vayda, 1973).

### 2.3 Wennberg’s “Small Area Variations”

In the late 1960s Vermont introduced a data system to monitor healthcare in the state. The design of this data system was such that one could analyze indicators of healthcare delivery at the community level. Capitalizing on the availability of this

---

\(^6\) Variation between: 1) regions within the state of Kansas (Lewis, 1969), 2) Canada and England/Wales (Vayda, 1973), and 3) regions within England and Wales (Feldstein, 1964).
data, and possibly spurred on by changes to the Social Security Act (and other legislation) introducing measures to regulate resource allocation and improve accountability, Wennberg and Gittelsohn (1973) undertook an accounting of healthcare resource inputs, utilization, and expenditures among 13 geographically defined healthcare service areas. What they found was substantial variation in the per capita 1) number of available beds, physicians, and hospital personnel, 2) number of hospital days, 3) rate of admission to nursing homes for individuals over age 65, 4) rate of discharge for various surgical procedures, and 5) hospital and nursing home expenditure. They defined these differences between communities as “small area variations”. Using age as a proxy for healthcare need, this study showed a poor correlation between community need (i.e. age of residents in the community) and the per capita number of physicians in the community. Based on this data, Wennberg and Gittelsohn’s stated that “variations in the health care experience of different Vermont populations may be explained more by behavioral and distributional differences than by differences in illness patterns” (Wennberg and Gittelsohn, 1973; p.1106). Furthermore, the authors state that such variation in use is an indication of uncertainty “about the effectiveness of different levels of aggregate, as well as specific kinds of, health services” (Wennberg and Gittelsohn, 1973; p.1107).

The recognition and reporting of small area variations by Wennberg and Gittelsohn is noteworthy for a number of reasons. First, their report generated interest in the study of practice variation. Whereas only a handful of studies
reporting or examining variation appeared in the literature before the year 1973, hundreds would appear over the next few decades after Wennberg and Gittelsohn’s work. For example, in a systematic review of practice variation, Paul-Shaheen and colleagues (1987) identified dozens of studies published in selected medical journals that either observed or discussed small area variations in just over a decade following 1973. Next, Wennberg and Gittelsohn offered an organized method by which to examine variation in healthcare service/resource use between populations in a (seemingly) policy relevant manner. Finally, their report emphasizes and perhaps offers evidence for the role of uncertainty (regarding the best course of action for treating many health conditions or disease states) as a mechanism by which variations in health service/resource use might be created.

In a follow-up paper, Wennberg expanded on his method for examining small area variations, providing some theoretical basis grounded in the economic behaviour of choice (Wennberg, Barnes, & Zubkoff, 1982). Like others, the framework proposed by Wennberg and his colleagues invokes the idea that, because of their specialized knowledge, physicians act on behalf of the patient in deciding the course of treatment. Furthermore, the small area method, in principle, accepts that variations in healthcare service use may arise between populations as a result of systematic differences in both incidence of disease and care seeking behaviour. Thus, under Wennberg’s (and colleagues’) framework, the general equation for estimating the probability of resource use for any individual within a population is determined by the product of 1) the probability of that individual suffering the
“disease” of interest, 2) the probability of that individual seeking physician care given that disease, 3) the probability of diagnosis given care is sought, and 4) the probability that the attending physician will use the resource of interest (e.g. surgery) given a positive diagnosis. The estimate of the rate of use for a population is a function of the sum of all individual use probabilities among patients within that population. In other words, “when the incidence of illness among populations is similar, when the members of the population contact their physicians in similar proportions, when the physicians make similar diagnostic and therapeutic decisions, and when available facilities are similar then the utilization rates among populations will be similar. When they are not similar, the differences can be explained by the differences among patients, physicians, available facilities, or all three” (Wennberg et al., 1982; p.812-3).

In concept, the framework provided by Wennberg is reasonably similar to the models for use provided by Ro (1969), and Andersen and Newmen (1973) in that they consider patient, physician, and environment level factors. However, Wennberg et al. (1982) are not specific with regards to which social factors should be accounted when determining rate of use, opting for the more general estimation of the probability that an individual will seek care. Compare this to Andersen and Newmen (1973) who provide a relatively extensive inventory of social factors that can prohibit or facilitate use. This difference can have important implications on how small area studies are constructed, analyzed, and interpreted. For example, consider two or more populations where individuals with a given disease seek care
in similar proportions. Under Wennberg’s framework, any variation in the rate of use among these populations would have to be attributed to systematic differences in the incidence of disease, actions of the attending physicians, or available resources. This overlooks the possibility that although two or more populations may be similar in the proportion of individuals with a disease seeking care, the proclivity towards accepting care among those within each population may be systematically different among the populations of interest. In other words, different populations may seek care in similar proportions yet for different reasons, and thus, may not necessarily want the same treatment. The models of Ro, and Andersen and Newmen are seemingly more sensitive to this possibility.

In application, the differences between the frameworks on healthcare service use are even more apparent. Wennberg’s experience with evaluating healthcare service use in Vermont offers a good example of how to apply the small area method, and how features of his theoretical approach/application of his framework impact the interpretation of the data in such a way that might have been different had the Andersen and Newmen model been applied. In practice, application of the small area method requires that geographic areas are defined and data is collected on the use of a particular set of healthcare services. In Wennberg’s example, the phenomenon of interest is the differential rate of surgical procedures among defined communities in Vermont. Based on the Wennberg’s framework, additional data regarding the burden of illness, the care seeking behaviour, and resource inputs must be collected and analyzed so that the sources of any residual variation
after adjustment for systematic population differences in these factors may be determined. When examining healthcare service use in Vermont, Wennberg would adjust for population differences in both the incidence/prevalence of disease (using age as a surrogate), and available resources (per capita beds, physicians, healthcare personnel). The result is a per capita rate of use of healthcare resources or services for each community that may or may not vary among them. Putting aside issues regarding the validity of age as a proxy for disease burden, or beds and healthcare personnel as a proxy for available resources, any residual variation in rate of use after adjustment for these factors must be attributed to either 1) differences between populations in care seeking behaviour, and/or 2) the physicians’ diagnostic ability and/or proclivity towards a certain kind of treatment action given diagnosis.

Here is where the differences in the framework for determining use have an impact. Wennberg and colleagues (1982) argue that the communities under investigation in Vermont were similar in their insurance, socio-economic, education, and race status. Therefore, any residual differences are attributed to the physicians. In contrast to the Andersen and Newmen (1973) model, numerous “predisposing” and “enabling” factors, including beliefs and values concerning illness and healthcare, occupation, and distance to healthcare services are not considered explicitly. Likewise, the small area method did not provide (or does not build in) a means by which to account for interactions between patients, physicians, and the environment as per the framework proposed by Ro (1969), including regulative and normative mechanisms that may differ between hospitals among the communities of interest.
These conceptual differences may not be of primary importance when comparing communities in 1960s Vermont, which at the time were quite likely to be homogenous with regards to values, beliefs, and socio-economic factors. The concern is with the approach more generally, especially when applied in communities where heterogeneity in social factors is likely.

2.4 “Uncertainty” in the Medical Evidence: The Underlying Reason for Variation?

Despite the potential limitations of his small area method in determining the sources of variation and their relative influence, Wennberg was able to provide both a reasonably good argument and empirical support for the role of uncertainty in the medical evidence as a mechanism for variation in medical/surgical practice (Wennberg et al. 1982). Wennberg and colleagues observed that for health conditions where there was less support in the medical literature as to the best course of treatment (i.e. more uncertainty), and where surgery was an option, there was relatively more variation between geographic areas in the rate of use of surgery or admission to hospital for those health conditions, and vice versa (Wennberg et al, 1982; Wennberg, 1984). However, it is not uncertainty itself, but rather the differential reaction to it that can cause variation. Patients, physicians, and healthcare managers all can potentially vary in how they react to uncertainty. Nevertheless, Wennberg would ascribe primary importance to the physician, further arguing that uncertainty and ambiguity in medical science creates
professional controversies as to the value of specific health services – where the physician stands in these controversies may dictate which therapy the patient receives (Wennberg, 1984).

The idea that uncertainty in medical/clinical science drives variations in practice has intuitive appeal. First, it preserves the image of the physician as competent. Where uncertainty exists, the physician makes treatment decisions thought to benefit the patient based on his or her clinical experience, interpretation of the available evidence, and knowledge of human physiology. Second, it provides a mechanism to reduce variation. That is, by eliminating uncertainty one can induce increased uniformity in the treatment patients with the same health condition receive. Furthermore, this view of “uncertainty” serves as a basis for the Evidence Based Medicine (EBM) paradigm, whereby clinical studies (including randomized controlled trials), systematic reviews, and subsequent clinical practice guidelines can, in theory, close knowledge gaps, and in turn, eliminate variation in practice. Heasman and Carstairs recognized this in response to their observation of variations in healthcare service use, when they stated, “our view is that, prima facie, the decisions which lie behind many of these very wide variations in practice are generally not founded on objective evidence, because this evidence is not available. Earlier one of us (Heasman, 1964) suggested that some of these questions could best be studied by the means of controlled clinical trials... Undoubtedly these trials offer the best solution, but they are both difficult and time-consuming” (Heasman and Carstairs, 1971; p.498).
Differential perception of or response to uncertainty in the medical evidence might explain variations in management decisions between individual physicians. For example, the response to uncertainty in one physician might be a proclivity towards the use of a therapy, diagnostic test, or surgical procedure, whereas in another an inclination to use such treatments or modalities sparingly. In other words, a lack of clear evidence as to the best course of management allows physicians to come to different conclusions regarding when to use a healthcare service. Considering variation at the regional level, it is unlikely that these “high” and “low” rate physicians are perfectly evenly distributed geographically\(^7\). Thus, it is reasonable that some regions will have a high number of “high” rate physicians, and other regions a high number of “low” rate physicians – the result is different rates in cumulative use between regions. In other words, small area variations might simply be due to chance. Eddy\(^8\) (1984) offers an alternative view as to how uncertainty causes small area variations. Eddy argues that if uncertainty caused physicians to practice at random, practice patterns would average out at the regional level (i.e. “high” and “low” use physicians would balance each other out both within and between regions). However, because working in an environment of uncertainty potentially exposes the physician to criticism about his or her management decisions, a safe and comfortable position is to adopt the management

\(^7\) In statistics we accept that some sample means will deviate from the population mean according to a “Normal” or “Gaussian” distribution. Thus, it is not unlikely that the rate of use in some regions will differ from others, strictly by chance.

\(^8\) Incidentally, David M. Eddy (1990) coined the term “Evidence-Based” in medicine, specifically in reference to practice policies.
practices of colleagues. Thus, Eddy contends that physicians will be inclined to “follow the pack”, and that variations between regions reflect differences in community standards regarding management of a specific medical condition (Eddy, 1984). Such community standards are maintained through normative mechanisms. However, there is limited empirical evidence to support the specific hypothesis that groups of physicians will adopt a similar practice style (although the influence of institutional restrictions and social context cannot be ruled out as important causal factors in this specific observation) (e.g. Westert, Nieboer, & Groenewegen, 1993).

2.5 “Practice Style”: The Physician as the Source of Variation

It is quite reasonable that uncertainty in the medical evidence creates opportunity for variation to develop. Certainly, if the best course of treatment for a given health condition were clear and obvious, one might expect deviation only where patient preferences or cost prohibit administering the evidence-based standard. However, such evidence rarely exists in medical practice – most management decisions are made under varying degrees of uncertainty – and although variation is more likely to emerge where uncertainty exists, uncertainty itself does not explain why one physician might adopt one style of practice and a different physician another. Wennberg offers perhaps a rather cynical view as to why this might occur, as is evident when he states that “most people view the medical care they receive as a necessity provided by doctors who adhere to scientific norms based on previously tested and proven treatments. When the
contents of the medical care “black box” are examined more closely, however, the
type of medical service provided is often found to be as strongly influenced by
subjective factors related to the attitudes of individual physicians as by science”
(Wennberg, 1984, p.7). These “subjective factors” manifest in what Wennberg
(1984) characterized as a physician’s “practice style”. Although “practice style” is
not explicitly defined (Stano, 1993), what is implied is that the unexplained
variation (i.e. that which is not explained by between populations differences in age,
and simple supply factors) in the use of healthcare services is a result of the
physician’s themselves, and not the patient or context. Indeed, Chassin9 (1993) has
argued that the “uncertainty hypothesis” is insufficient to explain observed
variations, but rather, they are a consequence of differing levels of “enthusiasm”
between individuals, or among groups of physicians regarding the usefulness of a
particular therapy or surgical intervention. While there are some data to support
this claim (Goel, Iron, & Williams, 1997; Wright et al., 1999), it is not clear why
physicians differ in their enthusiasm for one treatment strategy over another, nor
can one be sure that the observed enthusiasm is not itself a reaction to uncertainty.

9 Chassin demonstrates that surgeons in “high” and “low” rate regions select for patients in similar
ways, showing no significant differences in appropriateness, or in degree of surgical risk when
looking at practice related to carotid endarterectomy for transient ischemic attacks. The “high” rate
in some areas was explained by both a relatively higher number and proportion of surgeons
(resulting in a higher rate per capita) who perform carotid enarterectomy procedures. Furthermore,
“high” rate regions had proportionally more high volume surgeons when compared to “low” rate
regions. Chassin argued that this is “best characterized as enthusiasm for carotid endarterectomy”
among surgeons in “high” rate regions (Chassin, 1993, p.YS40). However, physician attitudes were
not measured in this study.
Regardless, “enthusiasm” implies that variations in the use of healthcare services are a direct result of (perhaps intrinsic) physician level characteristics.

Numerous “intrinsic” factors and/or personal characteristics have been shown to influence a physician’s management decisions. For example, Kostopoulou and Wildman (2004) found that when participant physicians reviewed case studies they differed in the information they gathered or requested, the importance they attached to such information, and how they interpreted it. Observed variations in their management decisions were attributed to those differences. Kalf and Spruijt-Metz (1996) found variations in management decisions, related to the degree of medical specialization. Tamayo-Sarver, Dawson, Cyduška, Wigton, and Baker (2004) found that emergency department physicians showed variability in how they responded to clinical information specified in patient vignettes when making decisions about prescribing opioid analgesics for pain. Survey data from the United States using patient vignettes showed that fear of a malpractice suit was a major determinant of a cardiologist’s propensity to test or treat (Lucas, Sirovich, Gallagher, Siewers, & Wennberg, 2010). This survey also demonstrated that many cardiologists reported they occasionally order tests to meet the expectations of the patients, referring physicians, or peers. Fear of malpractice was also a factor in determining the likelihood of referral among internists and family practice physicians working in managed care organizations (Franks, Williams, Zwanziger, Mooney, & Sorbero, 2000). Other factors shown to influence management decisions were clinical knowledge (experience, specialization, etc.), training environment, and
risk assessment and tolerance (Franks et al., 2000; McKinlay et al., 1998; O’Neill and Kuder, 2005; Reyna and Lloyd, 2006; van Miltenburg-van Zijl, Bossuyt, Nette, Simoons, & Taylor, 1997). However, apart from “fear of malpractice” (Lucas et al., 2010), studies have not examined the association of such factors with regional level variations.

Specific theories and/or mechanisms aside, non-experimental evidence demonstrating the relative importance of physician level factors in determining variation in healthcare service use is somewhat mixed. Physician level factors were indicated as the source of variation in at least four studies examining variation prior to the “discovery” of “small area variations”, as discussed earlier (Bunker, 1970; Freeborn et al. 1972; Childs and Hunter, 1972). Perhaps the best evidence in favour of physicians as the key source of variation comes from a study of primary physician services in Norway (Grytten and Sorensen, 2003). This study showed that physician-specific effects explained >50% of the variance in expenditures between physicians for a variety of procedures and laboratory tests. Furthermore, a physician's practice style typically remained stable in the event that he or she moved to a new municipality. Epstein and Nicholson (2009) found a similar result when examining management styles with regards to cesarean section. They observed that while obstetricians may learn from their colleagues, they are unlikely to significantly change their prior beliefs in when to use cesarean section. However, a study examining variation in length-of-stay for a variety of medical conditions demonstrated the opposite, with physicians adapting to the environment in which
they practice (de Jong, Westert, Lagoe, Groenewegen, 2006). That is, in the event that a physician worked at multiple hospitals, his or her management style related to length-of-stay was “contingent on” which hospital the physician was practicing in at the time (de Jong et al., 2006).

Physician level factors were considered the main source of variation between physicians in both how long they spent with patients during a visit, and tests they ordered for patients with arthritis, explaining >50% of the variance (Henke and Epstein, 1991). In addition, variations observed in studies by Feinglass, Martin, and Sen (1991), and Roos (1983) can be interpreted as due to idiosyncratic physician behaviour. Other studies show physician related factors to be of lesser importance compared to other factors, explaining only a small proportion of the total variance in management. This was the case for prescription rates (explaining only 3-10% of the variance), hospital admission rates (1%), ordering of lab tests (2-10%), length-of-stay (1-2%), overall resource use (2%), and overall lab costs (7%) (Davis, Gribben, Lay-Yee, & Scott, 2002; Hayward, Manning, McMahon, & Bernard, 1994; Hofer et al., 1999; Krein, Hofer, Kerr, & Hayward, 2002; Davis and Gribben, 1995).

“Practice style” implies that physicians will consistently apply their “rules” of practice when treating similar patients (Weinstein, Bronner, Morgan, & Wennberg, 2004). A study by Mousques, Renaud, and Scemama (2010) challenges this notion, where most of the variation (i.e. 73%) in antibiotic prescription for acute rhinopharyngitis was attributed to discrepancies within each physician’s practice (i.e. intra-physician variability). Finally, even when physicians differ in their
practice style, this alone does not necessarily explain variations in care across service areas, as was shown for admission rates for a variety of chronic conditions (Komaromy et al., 1996). The overall conclusion seems to be that while medical practice variations may in part be due to physician level factors or "practice style, the influence of such may not be equally important in all contexts. Furthermore, observed medical practice variations cannot be explained by physician level factors alone, intrinsic behavioural or otherwise.

2.6 Contextual/Environmental Sources of Variation

Characteristics specific to the context in which the care is delivered can influence management decisions. That is, physicians may be forced to make decisions regarding treatment within the confines of what the circumstances/environment allow. For example, Long (2001) argues in favour of a "resource constraint model", whereby physicians demand resources based on the clinical needs of the patient, modified by contextual constraints. These constraints are of three kinds: patient agency, organization, and environment (Long, 2001). Patient agency constraints include the patient’s health status or clinical needs, financial resources, availability, transportation limits, and demographics. The organization where the physician practices might institute mechanisms of control through fee schedules/capitation, resource use limitations, protocols or standards of practice, and or other normative or regulative means (organization constraints). Characteristics of the environment, including available beds, human resources, and
geography might also limit or dictate what care the patient might ultimately receive (environment constraints). Once all these accounted, any residual variation would be attributed to the “innate” characteristics of the physician. While Long (2001) does not deny that physician level factors are a source of variation, he believes that the relative influence of such factors is likely small.

There is empirical evidence to support the notion that variations in healthcare service use are not only due to physician level factors, but are more the result of differences in the availability of physical resources, such as beds, and medical equipment, and/or human resources, such as clinicians with specific expertise. These factors were implicated as the source of variation among many of the studies published in the years between Glover (1938) and Wennberg and Gittelsohn (1973), as described above. More recently, a study examining variations in healthcare service use among districts in the United Kingdom revealed an association between the rate of coronary artery revascularization procedures (i.e. bypass graft surgery and angioplasty) and differences in supply factors, such as the availability of a cardiologist, and proximity to the a regional revascularization centre (Black, Langham, & Petticrew, 1995). Likewise, the regional use of cardiac procedures in the United States of America was closely related to their availability (Pilote et al. 1995). Supply factors (in the form of bed availability) have also been implicated as the source of within centre variation in emergency department length-of-stay (Salas, Garrigo, Jimenez Vilchez, & Godoy, 2007).
As Long (2001) notes, contextual/environmental constraints need not be in the form of physical and human resource limitations. Localized social pressures and circumstances may also set the bounds within which physicians may act (Westert & Groenewegen 1999). This may be due to social pressures set by colleagues with whom the physician is mutually dependent, and/or regulatory mechanisms imposed by healthcare managers or funding agencies, such as insurance companies and governments (de Jong, Groenewegen, Spreeuwenberg, Schellevis, & Westert, 2010). Between-group differences in the opportunities and constraints (whether due to resource and/or social constraints) imposed by the work environment may result in geographic variations.

“Constraint centered” models of variation are consistent with Eddy’s (1984) assertion that where uncertainty exists groups of physicians will adopt similar practice. However, from this perspective the mechanisms for how a group of physicians adopt a common approach are extrinsic (i.e. group pressure, imposed regulation/resource restriction) rather than intrinsic (i.e. desire to be similar to group members) in nature. The implication of a “constraint centered” approach is that a greater proportion of the explained variance should be at the facility level than at the physician level. This was shown to be the case in a study focused on diabetes care (Krein et al. 2002). Furthermore, observations by de Jong et al. (2006) (described above) may also lend support to the “constraint centered” explanation. Overall, medical practice variations may in part be due to the practice environment in which the care is delivered. This can operate through two mechanisms: 1) social
pressures specific to each environment, and 2) differences in available resources within each environment.

2.7 Examining Variation – The General Picture and Questions Unanswered

The empirical evidence to date suggests three general sources of variation, operating at the level of the patient seeking care, the physician making management decisions, and the environment where care is delivered. The conceptual frameworks for examining variation proposed by Wennberg et al. (1982), and Long (2001) to some extent incorporate each of these general sources, as do the frameworks on healthcare services use suggested by Andersen and Newmen (1973), and Ro (1969). What is not clear is the relative influence of these sources, with Wennberg placing emphasis on the physician as the source of variation, whereas the others stress the importance of resource limitations or context (both patient and environment). This disagreement is reflected in the empirical evidence, where the proportion of explained variance attributed to each general source seems to vary among studies.

Although the general sources are considered in each of the frameworks described above, they are not comprehensive in identifying all the components of each source, nor are they necessarily explicit in describing how each should be measured or incorporated into studies examining variation. Even where the frameworks are explicit, some of the proposed components are often not directly measured (surrogate measures are used) or incorporated into the study design,
assuming that previous studies have even applied one of these frameworks\textsuperscript{10}. The consequence of incomplete identification and measurement of all the components is a significant proportion of the variance is left unexplained, as is the case in most (if not all) studies examining variation in healthcare service use. The unexplained variance is then typically attributed to one general source or another, likely depending on the preferences or values of the investigator, rather than defined by the conceptual framework (assuming the framework is correct). This problem in attributing variation to a particular source speaks to issues in both design (i.e. choosing the correct framework) and execution (i.e. correctly measuring all the components considered in the chosen framework).

Clinical scientists seem to give special attention to the physician as the primary source of variation. This is understandable, as physician driven variations in practice may be viewed as an issue with quality of care, rather than efficiency and equity in resource distribution. It is reasonable that clinical scientists might deem it most important that patients receive the most effective treatments, and leave issues of cost effectiveness and discrepancies in resource distribution to healthcare managers. However, a lack of theory and insufficient methods to examine variation make it difficult to fully understand the relative influence of each source and their components. Thus, it is not entirely clear whether meaningful variations indeed

\textsuperscript{10} Wennberg’s studies aside, the reviewed studies examining variation do not explicitly state a theoretical perspective or framework for investigation regarding use of health services, such as those identified here. Even among Wennberg’s studies, the reader must often assume the application of his framework outlined in Wennberg et al. (1982).
exist at the physician level after one considers all the patient and contextual/environmental factors. It is of particular importance that one determines the influence of the physician on creating observed variations, as substantial research and resources are put into reducing variation in medical practice at the physician level. For example, clinical trials and clinical practice guidelines exist for the purpose of reducing uncertainty as to the best course of treatment, with the hope that care will be standardized according to the best evidence. Likewise, physician “report cards” are developed to promote similar practice patterns among physicians. These activities, while resource intensive, will ultimately fail in reducing variation in a meaningful way if physicians and/or their reaction to uncertainty are not the reasons observed variations exist.

Variations research seeks to identify practice variations and determine their causes. However, much of the dialogue within and around these studies presupposes that variations are problematic, and thus, should be minimized (e.g. Wennberg, 2011). Unfortunately, it is not clear if it is the raw, unexplained, or (some or all of) the explained variation that should be minimized. It is reasonable to believe that few would argue in favour of minimizing the raw variation (i.e. before adjustment for any factors), as some of that variation might be appropriate, such as that due to differences in health status. Seeking to minimize only the unexplained variation may be equally problematic. As discussed above, a significant proportion of the variance in the variations studies remains unexplained. Again, it may be that we have not identified the appropriate factors to measure, and/or have the correct
methods for which to measure them. Blindly attempting to minimize variations that are not explained is unlikely to be effective, as it is not possible, by definition, to determine and ameliorate the cause. Indeed, it might result in poorer outcomes for patients if, for example, some of the unexplained variance were due to unmeasured differences in health status or patient preferences for care. Even if all the variance in any of the previous studies was explained, and all the sources, including their components and their relative influence were known, it is unclear as to which of these are problematic and why they might be.

Only those variations that can be explained but cannot be justified should be minimized. The described frameworks by Wennberg et al. (1982), and Long (2001) seem to indicate that it is only the residual variation, after all the factors outlined in their models have been considered, that is “unwarranted”. Wennberg would later argue that those variations that “cannot be explained on the basis of illness, patients’ preferences, or dictates of scientific medicine” are unwarranted (Mullan, 2004). Again, it is unclear if previous studies examining practice variations subscribe to one of these frameworks, or if they correctly measure the residual variation and/or account for those factors that might be considered warranted sources of variation. In other words, it is not apparent in most studies if and how they discriminate between those variations that should rightfully exist and those that are unwarranted. In the absence of a clear and accepted definition, any claim that observed variations might be “unwarranted”, may also be unwarranted. However, by what basis variations are or are not considered justified, and how this plays a
role in methods to examine practice variations require further examination.

Any variations that are considered unwarranted in a given population should be eliminated if one is to effectively use available resources to maximize the desired healthcare outcomes in that population. Those desired outcomes are determined by the principles of the healthcare policy under which the care is provided. While the literature on healthcare variation seems to generally oppose their existence, it does not explain what should be done beyond simply minimizing them, nor does it explain exactly how observed variations are relevant to the goals of healthcare decision makers, managers, and providers. The usefulness of variations studies in informing those stakeholders so they can make better decisions to meet policy goals requires further examination. Wennberg (2011) argues that information elucidated from variations studies should be integrated into decision making about healthcare service planning and resource distribution seemingly under the belief that doing so will help minimize unwarranted variations. However, this can only be the case if the metrics used in variations studies are aligned with the quality metrics defined by the health policy. Again, to what extent these metrics are aligned is not discussed in the common literature examining practice variations.

2.8 Preliminary Conclusions

Variation in the use of healthcare services has been a topic of interest in the mainstream healthcare literature for eighty years. This literature has been preoccupied with three issues: 1) identifying the existence of variations, 2)
determining their sources, and 3) minimizing them. One can argue that beyond identifying the general sources, the literature thus far has been inadequate in determining which specific factors cause variation, and their relative influence. Furthermore, without a clear distinction between those variations that promote suboptimal health outcomes or inefficient use of resources overall, and those that exist for good reason one cannot be sure that observed variations are at all meaningful (to healthcare stakeholders). Both of these issues are possibly the result of inadequate theory (or even use of established frameworks) regarding how to examine variation in the context of healthcare services. It is desirable that conclusions drawn from studies on practice variation be guided by established theory rather than the values of the individual researchers (as seems to be the case), which appears unlikely to occur until such time that a valid and functional theory of practice variation is established. This can, in turn, impact how the healthcare community engages observed variations, including the extent to which stakeholders desire the elimination of variation, how they devise and implement strategies to minimize them, and how effective they are in doing so. In this regard, this review of the historical literature on practice variations brings to light three features requiring further examination: 1) the extent to which physicians are the source of meaningful variation, especially where there is good scientific evidence of the relative benefits of available therapies, and patient and environmental characteristics are equal among physicians; 2) discriminating between variation that exists for good reason and that which might be considered unwarranted; 3) the
extent to which information reported in variations studies might be useful for healthcare planning.
Chapter 3: An Even Smaller Area Variation: Differing Practice Patterns Among Interventional Cardiologists Within A Single High Volume Tertiary Cardiac Centre

A common theme among studies of medical practice variation is the importance of physicians as a source of variation. Determining the relative influence of the physician requires that all other sources of variation, such as patient characteristics and resource availability, be held constant. Adjusting for these factors has been difficult due to data limitations in previous studies. The study presented in this chapter took a novel approach to determining if meaningful variation in exists between cardiologists by focusing on treatment recommendations among a group of cardiologists working in a single hospital, within a publicly funded healthcare system. Whereas previous studies used insurance or administrative databases, the existence of a clinical patient registry facilitated a relatively more detailed adjustment for potential patient population differences between cardiologists. This study suggests that physicians are likely a source of variation. However, the mechanisms by which physicians cause variation were not revealed.

An Even Smaller Area Variation: Differing Practice Patterns Among Interventional Cardiologists Within A Single High Volume Tertiary Cardiac Centre

Mathew Mercuri MSc\textsuperscript{1,3}, Madhu K. Natarajan MD MSc\textsuperscript{1,2,3}, Geoff Norman PhD\textsuperscript{3}, Amiram Gafni PhD\textsuperscript{2,3}

\textsuperscript{1}Heart Investigation Unit, Hamilton Health Sciences, Hamilton, Ontario, Canada;
\textsuperscript{2}Population Health Research Institute, Hamilton, Ontario, Canada;
\textsuperscript{3}McMaster University, Hamilton, Ontario, Canada

Corresponding Author:

Mathew Mercuri MSc,
Heart Investigation Unit, Hamilton General Hospital
237 Barton Street East,
Hamilton, Ontario, Canada, L8L 2X2,
Telephone: 905-527-4322 ext.44763;
e-mail: mercuri@hhsc.ca

Brief Title: Practice Variations Among Interventional Cardiologists

Word Count: 3677

Disclosures: None

Key Words: Practice Variation, Coronary Artery Disease, Practice Style, Single Center
3.1 Abstract

Background: Variations in the rate of use of common medical procedures/therapies are widely documented. Previous studies tend to focus on variations between either hospitals or geographic areas. Few studies examine within centre practice variations.

Objective: To examine if variation in treatment recommendations exist among highly trained interventional cardiologists (n=9) working in a single, highly collaborative tertiary care center.

Study Design and Setting: Data was collected from a local registry. A logistic regression model was used to estimate each physician’s odds of recommending revascularization therapy over medical therapy for patients with significant CAD. The analysis was repeated to estimate each physician’s odds of recommending percutaneous coronary intervention (PCI) over coronary artery bypass graft surgery (CABG) when the physician indicated the need for revascularization. Each physician’s odds were compared to those for a reference physician to yield odds ratios. The odds ratios were adjusted for multiple patient characteristics.

Results: The adjusted odds ratios of four physicians differed significantly from the reference physician (range: 0.8 to 2.9). Variation was also seen among physicians in the decision to recommend CABG rather than PCI once revascularization therapy was selected. The odds ratios ranged from 1.5 to 4.2.

Conclusion: Practice variations were seen despite case mix adjustment and similar resource and environmental constraints. The existence of within centre variations
may have implications on service delivery and planning. Research is needed to both identify the existence, and explain the determinants of "an even smaller area variation".

3.2 Background

Variations in the rate of use of common medical procedures/therapies have been widely documented [e.g. 1-10]. These studies suggest that the care an individual receives is contingent not only on his or her health state, but also on where and by whom the care is provided. Practice variations can have both negative and positive implications on healthcare, as was noted in a recent debate in The British Medical Journal [11]. Thus, it is important to better understand why variations exist, such that healthcare planning and patient outcomes can be optimized.

While previous studies have done much to improve our understanding of both the determinants of medical practice variations and the impact they may have on healthcare in general, these studies focus mainly on variations between either hospitals or geographic areas. However, practice variations may exist at lower levels of aggregation as well, such as in treatment decisions among physicians working within the same hospital. Their existence may provide additional service delivery challenges to local healthcare providers/managers. One cannot assume that the causes or the relative importance of variations at the regional or hospital level are the same at lower levels of aggregation. For example, consider one form of
medical practice variation identified by Wennberg as “variation in supply-sensitive care”, which arises from differences in the frequency in which individuals use healthcare services [12]. Wennberg has argued that this is in a large part determined by the amount of resources per capita allocated to the regional population/hospital. This factor may play a lesser role in determining practice variations between physicians within a single hospital, in the event that such variations exist. This may be especially true when determining the existence/causes of practice variations in a publicly funded healthcare environment where no difference exists between patients within a given hospital with regard to this aspect (e.g. hospital care in Ontario, Canada). Therefore, it is valuable to examine within centre practice variation as well, so that healthcare providers may better understand their causes (which may ultimately translate into more efficient strategies for dealing with unwarranted local variations). Only a few studies have examined variations among physicians within centres (e.g. 13-15). These studies have yielded mixed results regarding both the existence of practice variations within a group, and, in the event that they did exist, their determinants.

The purpose of the current study is to determine if meaningful practice variations exist among highly trained interventional cardiologists working within a university affiliated cardiac catheterization laboratory. Specifically, this study will focus on treatment recommendations (i.e. prescribed treatment rather than the treatment the patient ultimately receives) for patients diagnosed with significant coronary artery disease (CAD). Two treatment decisions will be assessed: a) the
decision to treat via revascularization versus medical therapy alone, and b) if
revascularization is prescribed, the decision to treat via surgery versus
percutaneous coronary intervention (PCI).

3.3 Methods

3.3.1 Study Site

This study was conducted at a large tertiary care centre in Ontario, Canada.
This center provides a full range of adult cardiac services, including diagnostics,
surgery (cardiac transplant excepted), interventions (arrhythmia, PCI, etc.), and
acute care. The cardiac catheterization unit is a high volume centre (>8000
procedures each year), servicing a population of 1.5 million people in a 7000-km²
area. While the unit provides a variety of services, the primary functions are
diagnosis via cardiac catheterization, and treatment of CAD via PCI (when
indicated). The unit consists of four cardiac catheterization suites and a 32-bed
reception/recovery area. This ensures an available bed for all patients requiring a
PCI. All four laboratories are adjacent to each other, and pairs of laboratories share
a common observation area. Services are publicly funded, ultimately operating
under the direction of the Ontario Ministry of Health and Long Term Care
(MOHLTC). The physicians are paid fees for service by the MOHLTC. Physicians are
not constrained by the hospital or the MOHLTC in their treatment
recommendations.
Referrals to the unit are highly selected. All patients must be referred by a cardiologist, or selected internists. All referrals then undergo screening by a regional cardiac care coordinator. While some cases are referred to specific physicians, most of the patients are referred to the group. These patients are subsequently assigned to a laboratory time slot based on catheterization laboratory availability, irrespective of the interventional cardiologists who is working that day. This is done at the discretion of the regional cardiac care coordinator. Physician schedules are not fixed in that the days each physician works vary from week to week. Further details regarding the cardiac catheterization unit in this study are provided elsewhere [16].

3.3.2 Population

This study focused on the practice of nine physicians working within a single cardiac centre. The nine physicians are all trained interventional cardiologists, having completed residencies in both internal medicine and general cardiology, and a fellowship in interventional cardiology. While the level of training is consistent among physicians, the location of such training, as well as the number of years of practice varies within the group. Many of the unit’s physicians have received additional graduate level training in health research methods and all participate as investigators in clinical trials. In addition, all physicians in this group are high volume operators, typically performing >500 cardiac catheterizations and >150 PCI procedures annually. These physicians performed catheterization laboratory
procedures exclusively at this centre during the observation period. Finally, all of the unit’s physicians are faculty members in the department of medicine at a hospital-affiliated university. Physicians who are currently training in the unit (i.e. residents and interventional cardiology fellows), or who are not trained in interventional cardiology were excluded from this study. This may be considered a highly collaborative group (e.g. they regularly discuss cases in a common viewing room, collaborate on clinical trials, and participate in a departmental journal club). Most notably, each physician regularly attends and performs case presentations at weekly departmental interventional cardiology rounds.

3.3.3. Data Collection and Analyses

This study utilized data collected from a local cardiac catheterization laboratory registry. This registry was implemented in 1997 for use in tracking demographics and clinical characteristics of patients referred to the catheterization laboratory. All data in this registry is prospectively collected. Details regarding the structure of the registry are described elsewhere [16].

Data regarding patients referred for a cardiac catheterization for suspected CAD between June 2006 and November 2008 were extracted from the registry. This period was chosen because during its duration there were no significant changes in practice guidelines for treating patients with CAD. Exclusions from the extracted sample were a) patients who previously underwent a coronary artery bypass graft surgery (CABG); b) patients presenting with an acute myocardial infarction (MI), as
such referrals are usually accepted on the intention to treat via PCI (and thus, decision to treat otherwise was post hoc based on clinical indication; e.g. emergency surgery required, contraindication to PCI, etc.) and c) left main disease > 50% as these referrals are far more likely to be referred to surgery. There is little expectation that there will be variation among physicians regarding treatment of both patients with acute MI and/or left main disease because there is substantial clinical evidence regarding benefit of one treatment over the others (as reflected in clinical practice guidelines; i.e. Class I recommendations) [17, 18].

The dependent variable in this study was the treatment prescribed by the interventional cardiologist after a diagnosis of significant CAD (defined below). Three treatment options were available: medical therapy only, referral for PCI, and referral for CABG. The interventional cardiologist indicated the prescribed treatment for each patient on a diagnostic case report form. This data was subsequently recorded in the registry. In a rare number of cases the physician indicated a referral to both PCI and CABG. Such cases were considered CABG referrals unless the patient received a PCI immediately following the diagnostic procedure (in which case, the treatment decision was considered PCI). This was considered a safe assumption, as 90% of all PCI procedures performed at this centre during the observation period were delivered in this fashion, and a decision to defer a PCI procedure (rather than perform it ad hoc) and refer to CABG likely meant the interventional cardiologist was seeking a surgical opinion regarding treatment.
A significant CAD diagnosis was considered any of the following (as defined by the treating interventional cardiologist): single vessel disease, two-vessel disease including the proximal left anterior descending artery (LAD), two-vessel disease including the distal LAD, two vessel disease not including the LAD, and three-vessel disease. A coronary artery (i.e. a vessel) was considered diseased if there was a greater than 70% narrowing within the vessel. A coronary artery was also considered diseased if there was a greater than 50% narrowing within a vessel when at least one other major vessel (i.e. LAD, circumflex, right coronary artery) contained a greater than 70% narrowing. All analyses excluded data from patients diagnosed with “mild” CAD (i.e. no vessel with a narrowing >70%), as all such patients were prescribed medical therapy alone.

A logistic regression model was used to estimate each physician’s odds of recommending revascularization therapy over medical therapy for patients with significant CAD. In order to do so, dummy variables were created for each physician. The physician with the most experience (in years) was used as the reference. A second logistic regression model was developed to estimate each physician’s odds of recommending PCI over CABG in the event that he/she indicated the need for revascularization therapy. Both models corrected for the following patient characteristics: age, sex, acuity on presentation (inpatient vs. outpatient), renal insufficiency (creatinine >200 µmol/L, not on dialysis), diabetes status, left ventricular ejection fraction (LVEF: <35% vs. ≥35%), presence of a cardiac surgical co-morbidity (at least one of: severe carotid stenosis (>70%), previous
stroke/transient ischemic attack, peripheral vascular disease (PVD), varicose vein, and/or chronic obstructive pulmonary disease (COPD)), and disease severity (5 levels; defined above). Both logistic regression analyses were performed using SPSS version 14. Microsoft Excel was used for all other analyses performed in this study. The risk of “over-fitting” the model was minimized, as each model had an observation-to-factor ratio of greater than 85.

3.4 Results

During the observation period, a total of 5458 patients were diagnosed with significant CAD by one of the nine interventional cardiologists. Medical therapy was recommended to 1029 patients, whereas a total of 4429 patients (81%) were referred for revascularization therapy (2869 for PCI; 1560 for CABG). Table 3-1 displays the clinical characteristics/demographics of patients in each of the three recommended treatment groups. The proportion of cases referred for revascularization varied among the physicians included in this study (range: 74-90%, goodness-of-fit p<0.01). Results from the logistic regression model demonstrated differences among the physician group in the odds of recommending revascularization therapy over medical therapy alone when compared to that of the reference physician, after adjustment for the factors outlined above. Five physicians did not vary from the reference physician. Among the physicians that did differ the most extreme (physician 4) was 2.9 times more likely to recommend revascularization when compared to the reference physician (Table 3-2). A Hosmer
and Lemeshow test yielded a p-value of 0.868, demonstrating a good fit of the data. However, these results should be taken with caution, as the classification table showed the model’s overwhelming inability to correctly predict patients receiving medical therapy.

Variation was also seen among physicians in the decision to refer to PCI rather than CABG once revascularization therapy was selected, with rates ranging from 59% to 73% (goodness-of-fit p<0.01), prior to adjustment. The logistic regression model examining the decision to recommend PCI over CABG performed much better than that studying the recommendation of revascularization therapy over medical therapy alone. Again, a Hosmer and Lemeshow test demonstrated a good fit of the data (p = 0.613). This model correctly predicted PCI cases (84%) and CABG cases (89%), for an overall correct classification of 87%. The model explained approximately 65% of the variance (Nagelkerke $R^2 = 0.645$). When compared to the reference physician, the odds of recommending PCI over CABG significantly differed for all but one physician (physician 6), after adjustment. One physician (physician 8) was 4.2 times more likely to recommend PCI over CABG to a patient when compared to the reference physician, after adjustment for the factors described above. Physician 8 was also significantly more likely to do so when compared to physician 6, although not to the same extent in magnitude as when compared to the reference physician. The estimated odds ratios for each physician as estimated by the PCI vs. CABG model are displayed in Table 3-2.
3.5 Discussion

This study demonstrates variations in treatment recommendations among interventional cardiologists working within the same cardiac unit. That is, there was up to a 4-fold difference between physicians in the odds of choosing one therapy over another, after adjustment for case-mix differences. Such treatment decisions are not trivial as there are important differences in both the invasiveness and cost of each treatment, and the nature of the commitment by the patient. Our study differs from previous studies examining within centre variation in that such studies focused on aspects of practice that are not completely under the physician’s control (e.g. length of stay) (13-15), whereas our study focused on the treatment recommendations by the physicians. Thus, the results from the study presented here are not directly comparable to those of previous studies.

Numerous arguments have been put forward to explain why variations in medical practice exist. One of the most compelling arguments maintains that practice variations are a result of differences in resource and/or environmental constraints (12, 13, 19, 20). Such factors are unlikely to play a significant role in explaining the variation observed here, as this study examined variation in practice within a single hospital unit, and thus, each physician was influenced by the same organizational/environmental constraints. Differences in case mix could play a role as well. However, previous studies have consistently shown that case mix differences alone do not sufficiently account for observed variations in medical practice (21-28). It is unlikely that case mix differences between physicians explain
the variation seen in this study given both the adjustment for multiple clinical characteristics and demographics, and how the majority of cases were assigned to each physician (i.e. based on laboratory availability, irrespective of the operator).

It is possible that the observed differences in treatment practices among the physicians in this study are due to factors associated with the physicians themselves. Wennberg and others have argued that individual physicians develop a preference for particular therapies, possibly due to uncertainty regarding the benefit of available treatments, and/or enthusiasm for particular technologies/therapies [29-33]. We know that from laboratory-based experiments, such as vignette-based research investigating medical decision-making, variations in a variety of physician-intrinsic factors may account for differences in medical decisions between operators. Such factors include the physician’s age, specialty, clinical knowledge, training environment, tolerance of risk, and how the physician interprets the available evidence and/or clinical information (34-43). Given that resource/environment and case-mix factors appear to be insufficient in explaining the variations seen in this study, it is not unreasonable to believe that physician related factors, such as those described above, are playing role.

Unfortunately, one cannot determine the exact mechanism(s) based on the data available in the registry. There was no correlation observed between years of experience and propensity to recommend revascularization over medical therapy alone, or PCI over CABG. Furthermore, factors such as sex and medical specialty cannot explain the variation observed here, as all the physicians in this study were
male and of the same specialty (i.e. interventional cardiology). However, further examination of the data did provide insight into the possible role of uncertainty among operators as a determinant of practice variation. For example, the treatment benefit derived from one revascularization therapy over the other appears to be unclear for patients with two-vessel disease [44]. This uncertainty is reflected in both the surgical and interventional cardiology practice guidelines [17, 18]. Therefore, in the event that operator uncertainty was a determinant of practice variation, one would expect the variation to be most pronounced when focusing on patients with two-vessel disease. This was indeed the case. A repeat of the regression analysis after excluding patients with single or three-vessel disease revealed a 6-fold difference between physicians in the odds of recommending PCI over CABG. Interestingly, this observation appears to counter the argument by Eddy (32), which maintains that physicians will adopt a common practice in situations where there is uncertainty regarding the best course of action. The physicians in this study have not adopted a common practice to treating two-vessel disease. Rather, their practice seems to reflect ambiguity in the practice guidelines. This is especially surprising considering the highly collaborative nature of the group. While this study seems to support the general argument that “practice style” is a determinant of medical practice variation, it should be noted that this study was not specifically designed to examine such factors. Therefore, the explanations provided above should be taken with caution and should be considered as hypothesis generating.
This study demonstrates that practice variations are possible even among similarly trained physicians working in the same department. Recognition of this fact may have implications on both health services research and health policy. First, variations such as those seen here may make it difficult for healthcare managers to plan and deliver services at a local level. This may prompt managers to institute regulative mechanisms in order to minimize variation. In some circumstances, managers may even be tempted to encourage the adoption of the “least expensive” practice. Such practices may be unreasonable without a thorough understanding of why such variations exist. It is possible that these variations are completely justified. Second, strategies designed to reduce unwarranted variations may not be effective if they target the wrong mechanisms. We cannot assume that these mechanisms operate equally at both the regional and local level. It would seem from this study that local variations are more the result of physician related factors rather than supply factors, which are likely to play a larger role at the regional/inter-hospital level. Therefore, in the event that strategies to reduce unwarranted variation at the regional level were based on the understanding that supply factors are the key determinant, such strategies would have minimal impact on reducing variation at the local level. Finally, the traditional approach among researchers and policy makers to focus on variations at the regional level may miss important variations within groups. Micro level variations may “wash out” at higher levels of aggregation. This may potentially give those monitoring practice patterns a false sense of the magnitude of treatment variations among individual patients.
Based on these arguments, it may not be prudent for healthcare managers and health services researchers to focus exclusively on variations at higher levels of aggregation. Expanding the investigation to lower levels of aggregation will allow for the identification of potentially unwarranted practice variations and may also yield valuable information for improving our understanding of their causes.

3.5.1 Limitations

There are several limitations typical of a retrospective observational design. Most notably was the influence of unknown/unmeasured factors on variation. Some continuous variables were regrettably dichotomized (i.e. LVEF and renal insufficiency/creatinine level) due to how they are recorded in the database. However, this was not seen as a major limitation, as this is somewhat inline with how these variables are viewed when making decisions about treatment (i.e. good vs. poor LVEF as opposed to 35% vs. 42%). In addition, it could be argued that the disease categories used in this study were not specific enough. For example, not all “one-vessel disease” diagnoses are equal. Both the location of the narrowing within the vessel and lesion complexity are important factors when prescribing treatment. It was assumed that the distribution of such factors within each category was reasonably equal between physicians. Unfortunately, the sample size limited further subdivision of categories, as this would result in an insufficient number of cases for analyses within each group. Furthermore, data regarding lesion length and diameter was only available on a subset of patients in the database (i.e. those who
underwent PCI). Despite these limitations, the case-mix adjustment in this study was more comprehensive than many previous studies examining practice variations, where case-mix adjustment often consisted of nothing more than age and sex. Regrettably, we were unable to collect data regarding patient preference. Thus, we cannot assess the influence of patient preference on between physician variations. It is possible that patients with a certain preference sought out specific physicians. Likewise, some physicians may be more sensitive to teasing out patient preferences. Such information could provide valuable insight into the causes of physician style.

Methodologies that employ data from registries/administrative databases can often identify practice variations. Unfortunately, they seem to be inadequate for identifying the specific mechanisms responsible for the observed variation. Studies that utilize medical vignettes have been shown to be useful in identifying such factors. However, vignettes focus on the few factors that can be described in text by the investigator and may potentially “miss” some of the nuances in the patient condition/preference. Studies such as the one presented here can compliment information extracted from vignette-based research. Regrettably, we could not make use of patient vignettes in the current study because 1) our sample size was inadequate to correlate variations in physician related factors to the observed practice variations, and 2) at the time this study was conducted, some of the physicians in this study were not available for data collection beyond what was available in the registry. Therefore, future studies should employ both techniques.
to examine physician related causes of practice variation. This would include further investigation into identifying if there exist variations in how physicians interpret evidence and/or the patient’s health state. It may also be interesting to see if physicians differ in the shared decision making process with patients. In the event that environmental/supply factors and case mix are insufficient to explain practice variations, the best we can do is to attribute the observed variation to a vague notion of “practice style”. Research is needed to both identify the existence, and explain the determinants “an even smaller area variation”.
3.6 References


<table>
<thead>
<tr>
<th></th>
<th>Med Rx</th>
<th>PCI</th>
<th>CABG</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>n</td>
<td>1029</td>
<td></td>
<td>2869</td>
<td></td>
</tr>
<tr>
<td>inpatient</td>
<td>529</td>
<td>51.4</td>
<td>1947</td>
<td>67.9</td>
</tr>
<tr>
<td>female</td>
<td>358</td>
<td>34.8</td>
<td>943</td>
<td>32.9</td>
</tr>
<tr>
<td>renal insufficiency</td>
<td>21</td>
<td>2.0</td>
<td>32</td>
<td>1.1</td>
</tr>
<tr>
<td>diabetes</td>
<td>346</td>
<td>33.6</td>
<td>704</td>
<td>24.5</td>
</tr>
<tr>
<td>LVEF &lt;35%</td>
<td>113</td>
<td>11.0</td>
<td>126</td>
<td>4.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Med Rx</th>
<th>PCI</th>
<th>CABG</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>&gt;=1 Sx Comorbidity</td>
<td>239</td>
<td>23.2</td>
<td>388</td>
<td>13.5</td>
</tr>
<tr>
<td>severe carotid stenosis (&gt;70%)</td>
<td>26</td>
<td>2.5</td>
<td>47</td>
<td>1.6</td>
</tr>
<tr>
<td>previous stroke</td>
<td>90</td>
<td>8.7</td>
<td>169</td>
<td>5.9</td>
</tr>
<tr>
<td>varicose vein</td>
<td>27</td>
<td>2.6</td>
<td>52</td>
<td>1.8</td>
</tr>
<tr>
<td>COPD</td>
<td>47</td>
<td>4.6</td>
<td>81</td>
<td>2.8</td>
</tr>
<tr>
<td>PVD</td>
<td>128</td>
<td>12.4</td>
<td>142</td>
<td>4.9</td>
</tr>
<tr>
<td>disease severity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1VD</td>
<td>581</td>
<td>56.5</td>
<td>1864</td>
<td>65.0</td>
</tr>
<tr>
<td>2VD incl prox LAD</td>
<td>112</td>
<td>10.9</td>
<td>284</td>
<td>9.9</td>
</tr>
<tr>
<td>2VD incl distal LAD</td>
<td>107</td>
<td>10.4</td>
<td>286</td>
<td>10.0</td>
</tr>
<tr>
<td>2VD w/o LAD</td>
<td>40</td>
<td>3.9</td>
<td>177</td>
<td>6.2</td>
</tr>
<tr>
<td>3VD</td>
<td>189</td>
<td>18.4</td>
<td>258</td>
<td>9.0</td>
</tr>
</tbody>
</table>

Table 3-1: Clinical characteristics/demographics of patients in each of the three recommended treatment groups. VD: Vessel Disease.

<table>
<thead>
<tr>
<th>Physician</th>
<th>Revascularization vs. Medical Therapy</th>
<th>PCI vs. CABG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds ratio</td>
<td>95% Confidence Interval</td>
</tr>
<tr>
<td>2</td>
<td>0.99</td>
<td>0.71-1.39</td>
</tr>
<tr>
<td>3</td>
<td>1.92</td>
<td>1.31-2.81</td>
</tr>
<tr>
<td>4</td>
<td>2.91</td>
<td>2.00-4.24</td>
</tr>
<tr>
<td>5</td>
<td>1.14</td>
<td>0.81-1.60</td>
</tr>
<tr>
<td>6</td>
<td>1.86</td>
<td>1.32-2.61</td>
</tr>
<tr>
<td>7</td>
<td>0.77</td>
<td>0.57-1.05</td>
</tr>
<tr>
<td>8</td>
<td>1.17</td>
<td>0.83-1.65</td>
</tr>
<tr>
<td>9</td>
<td>1.14</td>
<td>0.79-1.65</td>
</tr>
</tbody>
</table>

Table 3-2: Adjusted odds of prescribing revascularization over medical therapy, and PCI over CABG for each physician in the study compared to that for the reference physician.
Chapter 4: The Influence of Patient Social Context in Determining Emergency Medicine Physicians’ Management Decisions Relative to Clinical Practice Guidelines

Although the study presented in the previous chapter demonstrated variation between physicians after adjustment for resource availability and clinical characteristics of the patient population, the mechanism(s) by which this variation arises could not be determined. The study presented in this chapter was designed to examine the role of one possible mechanism – the patient’s social context. Despite the existence of evidence-based clinical practice guidelines to assist the physician in determining how care should be managed, participants in this study often deviated from those guidelines when information pertaining to the patient’s context was presented. This study focused on emergency medicine specialists rather than cardiologists because of sample size requirements and the need for a resident group. What is presented here should be considered a pilot study.

4.1 Introduction

Previous research indicates that physicians differ in the rate at which their patients use healthcare services, despite similar health status (e.g. Mercuri, Natarajan, Norman, & Gafni, 2012). These observed variations in the rate of use seem to be widest where there is professional uncertainty as to the best course of action given a specific medical condition, likely due to a lack of clear evidence in the
medical literature (Mercuri et al., 2012; Wennberg et al., 1982; Wennberg, 1984). Some argue that where such uncertainty abounds, management decisions are a result of each attending physician’s (differing) reaction to this professional uncertainty, as much influenced by subjective, physician related factors related as they are by scientific principles (Wennberg, 1984; Eddy, 1984).

Clinical practice guidelines (CPGs) were introduced as a means to reduce professional uncertainty as to the best course of treatment for a given medical condition. Such guidelines are developed through review and synthesis of the best available scientific evidence, and consensus by a panel of medical and scientific experts. CPGs are intended to promote standardization of medical practice along the lines of scientific, or best (available) evidence of effectiveness principles. Thus, by reducing professional uncertainty, CPGs might operate to reduce variation in medical practice (including the rate of use of healthcare services) (Davis and Taylor-Vaisey, 1997; Woolf, Grol, Hutchinson, Eccles, Grimshaw, 1999).

While medical practice is grounded in clinical science, healthcare management decisions are based not only on what the medical evidence determines will most likely result in the desired medical outcome, but also on what the individual patient will or can tolerate given his or her specific context (assuming resources required to provide the CPG recommended care are available). For example, Andersen and Newman (1973) outline a number of patient related factors that can influence the use of healthcare services, including affordability, access, and attitudes (both personal and family) towards health and healthcare. Likewise, Ro
(1969) argues that what the physician recommends is mediated in part by their response to each patient’s “choice conditioning” factors, which may be personal or situational in nature. Thus, even when professional uncertainty is minimized (e.g. when evidence based CPGs are available), management decisions may be determined by contextual factors beyond diagnosed health status. Previous studies have examined why physicians do not always adhere to CPGs (see the review by Cabana et al., 1999). These studies primarily focus on deficits in physicians’ knowledge or awareness of CPGs, with no explicit examination of social context beyond patient preference.

The purpose of this study was to determine if factors related to an individual patients’ social context are associated with deviations in healthcare management from that recommended by CPGs. Social context factors of interest include patient’s occupation, addiction history, proximity to care, expectations of treatment, and factors related to home life (e.g. lives alone). This study focused on point of care management decisions for common emergency department (ED) cases where established CPGs exist. The primary goal was to test the following hypothesis:

1) Given a case where an established CPG exists, the rate at which experienced emergency medicine (EM) physicians’ management decisions will reflect those recommended by the CPG will be lower when a factor related to the social context of the patient is present.
Whereas information regarding the most effective treatment for a given medical condition might be acquired through study of the medical literature, effectively tailoring care to a given patient’s situation is likely learned through experience with patients. Thus, one might reasonably expect that more experienced physicians are more adept at integrating the patient’s social context into their management decisions. As a subsequent focus, this study tests the following hypothesis:

2) Compared to residents, experienced EM physicians’ management decisions will less often reflect CPG recommendations when a factor related to the social context of the patient is present.

4.2 Methods

4.2.1 Participants

Participants were recruited from among both experienced and resident emergency medicine physicians practicing in the Hamilton Ontario region. Experienced physicians were defined as those having Royal College of Physicians and Surgeons certification in Emergency Medicine, and at least 5 years of practice in EM. Residents were defined as those enrolled in a residency program in EM. Recruitment of residents was restricted to those in their first 3 years of residency.
4.2.2 Survey Instrument

Twelve pairs of clinical scenarios outlining typical cases EM specialists might encounter in practice were developed in consultation with an experienced and certified EM physician. One case in each pair included a “social variable” (SV) that described the patient’s context. Cases were chosen from among four domains (two related to diagnostic testing and two related to treatment). Specifically, cases focused on 1) ordering radiography for suspected ankle injury, 2) ordering computed tomography (CT) for suspected head injury, 3) prescribing antibiotics for sore throat (i.e. suspected streptococcal pharyngitis), and 4) prescribing warfarin for atrial fibrillation, where the patient has been referred to a cardiologist. The basic cases were developed to meet the terms of one of the following CPGs: Ottawa Ankle Rule (Stiell et al., 1994); Canadian CT Head Rule (Stiell et al., 2001); Centor Score/McIssac Score for Strep Pharyngitis (Centor, Witherspoon, Dalton, Brody, & Link, 1981; McIsaac, Goel, To, & Low, 2000); Atrial Fibrilation CHADS2 Score (Gage et al., 2001). Each case was then paired with one of identical content with added information regarding the “social variable”. Three pairs of cases were developed for each domain. Two additional experienced and certified EM physicians reviewed the cases to assess content and face validity. A brief summary of the cases and their associated SV are presented in Table 4-1.

Two surveys were developed based on these cases. Each survey contained one case from each pair (a total of 12 cases per survey). The surveys were balanced so that half the cases contained the SV; with each survey a mirror image of the other.
Prior to the presented cases, the survey included questions regarding both participant demographics, and characteristics of practice environment (described below). Finally, the survey ended with four questions regarding the participants’ familiarity with each of the above-mentioned CPGs.

4.2.3 Data Acquisition

This study used a web-based survey design administered via LimeSurvey. Participants were randomized to one of the two survey instruments, stratified according to experience (resident vs. certified EM physician). Once randomized, a unique survey link was prepared and forwarded to each participant via email. Only a single survey attempt was possible for each link, and the survey needed to be completed in a single session. A reminder email was sent out one week after the initial invitation to encourage participation.

Data regarding the participant’s demographics and work environment were collected as follows: 1) gender (male/female), 2) experience (resident/EM specialist), 3) number of years since completion of EM residency, 4) location of practice (rural/urban), 5) type of hospital of primary practice (academic/community), 6) approximate number of cases per shift, 7) approximate number of shifts per month, 8) monthly department patient volume. For each presented case, the participant was asked to indicate if he/she would order the diagnostic test or prescribe the treatment in question (yes/no). The participant’s familiarity with each CPG also collected (yes/no).
Participants received a small honorarium for participation in the study. The dataset was stripped of personal identifiers prior to analyses. This study received approval by the McMaster University/Hamilton Health Sciences research ethics board.

4.2.4 Analysis

The odds of following the CPG recommendation were calculated in both the “SV”, and “no SV” conditions, stratified by expertise (residents vs. experienced). A chi-square test was used to examine if the odds ratio was significant when comparing the odds of following the CPG when the SV was present versus that when the SV was not present. The average number of times the participant’s recommendation followed the CPG was calculated in each condition, again stratified according to experience. A paired t-test was used to determine if the rate at which CPGs were followed differed between conditions. Finally, a chi-square test was performed to determine if the odds of following the CPG in each condition differed according to expertise. A p-value of 0.05 was considered significant. All analyses were performed using Microsoft Excel or SPSS version 20 for Mac OS.

4.3 Results

Requests for participation were sent out to 30 individuals, evenly divided between experienced EM physicians and EM residents. Data was collected from among 10 experienced physicians, and 12 residents in January of 2013. Five EM
physicians and 7 EM residents completed survey form 1, whereas the balance of participants completed survey form 2. All of the participants indicated familiarity with all of the above-mentioned CPGs. All of the participants indicated as working primarily at an urban teaching hospital. Participant demographics and characteristics of work environment are described in Table 4-2.

Experienced physicians were significantly less likely to follow CPG recommendations in the presence of a SV compared to when the SV was not present (53% vs. 77%; OR=0.35, 95%CI: 0.16-0.76, p=0.007). This finding supports our primary hypothesis. The rate at which residents would follow CPG recommendations did not significantly differ between conditions (67% for SV condition vs. 75% for no SV; OR=0.67, 95%CI: 0.32-1.38, p=0.27). Experienced physicians on average followed the CPG recommendation in 1.4 fewer cases when the SV was present, compared to when it was not (95%CI: 0.09-2.71, p=0.04), whereas residents did so for an average 0.5 fewer cases (95%CI: -0.29-1.29, p=0.19). Only 3 of 10 experienced physicians followed the CPG recommendation at an equal or greater rate in the SV condition compared to the no SV condition, whereas this was the case in 6 of 12 residents.

We found no difference in the rate at which experienced physicians followed CPG recommendations in the no SV condition, when compared to residents (77% vs. 75%; OR=1.1, 95%CI: 0.49-2.44, p=0.82). While the rate at which management recommendations reflected CPGs in the SV condition was lower for experienced physicians compared to residents (53% vs. 67%), the odds of following the CPGs in
this condition did not significantly differ between groups (OR=0.57, 95%CI: 0.28-1.16, p=0.15). Figure 4-1 shows the rate of adherence to CPGs for both the experienced and resident physicians in both the SV and no SV conditions.

Complete agreement was achieved among experienced physicians in only 5 of 12 cases in the no-SV condition, and only 3 of 12 cases in the SV condition, with more than one physician differing from the group in the majority of cases. Furthermore, among the 5 cases where there was complete agreement, variation in the management decision was observed in 3 of those cases when the SV was introduced. Similar results were observed among residents (i.e. agreement in 4 for no-SV, and 2 for SV; variation in 3 of 4 cases when SV introduced).

4.4 Discussion

Experienced physicians’ management decisions were less likely to adhere to CPGs when they were presented case studies that included information on the patient’s social context. This observation supports our primary hypothesis. However, the data did not support our secondary hypothesis that experienced physicians are less likely to follow CPGs in the SV condition, when compared to residents.

While the observed management decisions of those physicians participating in this study typically reflected published CPG recommendations, adherence to those CPGs was incomplete. Numerous studies have investigated the reasons for poor adherence by physicians to CPGs (see Cabana et al., 1999). The majority of
these studies focused on lack of awareness or knowledge of current CPGs, and lack of belief or trust in published CPGs as the primary culprits (see Cabana et al., 1999). For example, in a systematic review, Choudhry, Fletcher, and Soumerai (2005) observed an inverse relationship between physician experience and quality of care (most often defined as CPG adherence in their study), which the authors attributed to a lack of sufficient “factual knowledge” (p.269) among experienced physicians. In contrast, Persell et al. (2010) determined that many of the deviations in practice from CPGs could be justified through acceptable “medical exceptions” related to the patient’s health status. Likewise, Hughes, McMurd, and Guthrie (2013) and Boyd et al. (2005) argue that CPGs are in some cases inappropriate for individuals with multimorbidity, as strict adherence to multiple guidelines would result in an unreasonable treatment burden – a feature that is not consistently accounted for in CPGs. Fewer studies investigated the potential role of “external barriers”, such as resource availability and/or patient preference (see Cabana et al., 1999). While many acknowledge patient preference as a warranted reason for non-adherence, the role of patient values and preferences in determining how care will be managed receives little attention among prominent CPGs (e.g. McCormack and Loewen, 2007). Furthermore, even where patient values and preferences are acknowledged, it is not clear as to how they are to be integrated with “evidence” of benefit when determining how care will be managed.

Our findings do not support the notion of some that adherence is related primarily to physicians’ knowledge of CPGs (e.g. Choudry et al., 2005). All of the
participants indicated familiarity with the case related CPGs, and management
decisions reflected such CPGs in the majority of cases in the control (i.e. “no SV”) condition, irrespective of EM experience. Despite this demonstrated knowledge, management decisions deviated from CPGs more often in the presence of a contextual factor (i.e. “SV”), especially among experienced EM physicians. These observations imply that the “lack of adherence” among experienced physicians may be less about assumptions regarding their level of CPG knowledge, and more about their sensitivity to each patient’s unique needs. Perhaps given opportunity to manage the care of many individuals, experienced physicians have developed an appreciation for how evidence-based therapies or diagnostic tests are tolerated or provide benefit in the context of both the patient’s health and social states. Furthermore, relevant aspects of the social context may not entirely be captured by patient preferences and values. Rather, management decisions, while based on the best available evidence, are tailored according to the physician’s perception of potential benefit given the whole experience of the patient, of which preferences and values are a part - as are factors related to the patient’s health status (co-morbidity, severity of disease, etc.), ability to adhere to treatment recommendations, ability to recover from surgery, and burden the therapy might impose on the individual/family members, among others. Social variables as conceptualized in this study cannot be viewed strictly as potential “medical exceptions”; rather, they may indicate a different class of “exceptions” to CPGs that
may require consideration when assessing quality of care based on physician adherence to guidelines.

Best practices for CPG development endorse the use of the best available evidence from the scientific literature (Brouwers et al., 2009; The AGREE Collaboration 2003). Commonly used tools to assess the quality of evidence, such as GRADE (GRADE Working Group 2004), give emphasis to evidence derived from randomized controlled trials (RCT). However, RCT’s devalue potentially important aspects of the individual that cannot be (easily) quantified (Tonelli, 1998).

Advocates of CPGs acknowledge the potential tension between clinical trial evidence regarding the best course of therapy for a given health state and how to manage care given the patient’s context. The Canadian Medical Association Handbook on Clinical Practice Guidelines states, “that CPGs are not intended to provide guidance in all circumstances and for all patients” (Davis, Goldman, & Palda, 2007, p.3), as the authors recognize that “their more general nature renders them insensitive to the particular circumstances of individual cases” (Davis et al., 2007, p.3). Likewise, Woolf et al. (1999) raise concern that “the frequently touted benefit of clinical guidelines – more consistent practice patterns and reduced variation – may come at the expense of reducing individualized care for patients with special needs” (p.529).

In some cases, consideration of the individual patient’s circumstance or “special needs” has been incorporated into CPGs. For example, the decision to prescribe highly active antiretroviral therapy (HAART) in HIV-infected individuals must balance the potential benefits of therapy with the risk of drug resistance as a result
of non-adherence by the individual (Wainberg and Friedland, 1998; Bogart, Catz, Kelly, Benotsch, 2001). Thus, a consideration of an individual’s non-adherence potential is incorporated into the CPGs for HAART (Carpenter et al., 1998). However, the extent to which other CPGs consider adherence and/or other contextual or “non-medical” factors that might affect care, or how they should be integrated with “medical” factors/trial evidence into management decisions is not clearly described in the literature.

Provided the goal of a healthcare system is to maximize health outcomes (i.e. mortality) to the population as a whole, RCT driven management decisions, and by extension CPGs, might be useful at the population level. Management decisions, on the other hand, are made at the individual level. Tonelli (1998) notes that at the individual level “illness is inseparable from other aspects of existence” (p.1237). This position is even supported by key advocates of the evidence-based medicine paradigm for management of care. For example, when determining if CPG recommendations are applicable to a patient, Wilson, Hayward, Tunis, Bass, and Guyatt (1995) suggest, “You should look for information that must be obtained from and provided to patients for patient preferences that should be considered. It is important to consider whether the values assigned (implicitly or explicitly) to outcomes could differ enough from your patients’ preferences to change a decision about whether to adopt a recommendation” (p.1632). This position is consistent with the presented findings, where physicians were influenced by patient context.
Although this study was not designed to examine if CPGs reduce variation, our results imply that they do not eliminate it. Furthermore, it appears that the patient’s social context might promote variation among both experienced and resident physicians. Thus, variation (beyond that due to differences in health status) may not be strictly due to reaction to “uncertainty” regarding the best course of action, or some physicians ignoring the “evidence”, but also may be the result of systematic differences between populations regarding the social context of patients, and/or how physicians might perceive its importance when determining the best course of care. However, one unique aspect of this study is that it goes beyond the usual model of “environment/physician/patient” as the sources of variation. Regardless of how we defined the SV, it does not strictly “fit” into a “patient characteristic”. Rather, this study points out that variation may be due to the interaction between the physician and patient, as in this experiment it was the physicians’ interpretation of the importance of the SV that influenced the decision.

This study could not determine precisely why some physicians’ management decisions differ from CPG recommendations in some cases and not others (irrespective of the presence of social context), nor could it determine if those deviations we attributed to social context would result in better patient outcomes. However, if one considers social context to be a justifiable reason to deviate from CPGs this would have significant implications on how we examine variation. Where there is good evidence of benefit for a particular therapy or diagnostic test given a particular health status, any differences between populations that cannot be
accounted for by differences in health status might be considered “unwarranted”. In the event that the populations also differed in “social context” independent of health status, one might incorrectly consider the residual variation “unwarranted” unless the “social context” were measured and incorporated into the analyses (Mercuri and Gafni, 2011). Consideration of “social context” would mark a shift in the approach to quantifying variation, where previous research primarily accounted for population differences in demographics and in some cases disease severity (Mercuri, Birch, & Gafni, 2013). Furthermore, CPGs might imply that there is a “right” rate of use that should not differ between populations after accounting for health status. The data presented here might bring into question the concept of a universal “right” rate of use; rather, what is considered correct in each population may depend on local characteristics related to the “social context” of that population (Mercuri et al., 2013).

4.4.1 Limitations

This study has several limitations that are typical of experimental designs. Most notably is the extent to which results based on hypothetical cases are generalizable to actual practice. In addition, the presented results are based on management decisions among a small group of EM physicians in a single, albeit large urban community. Thus, the idea that patient “social context” can affect adherence to CPGs should be considered preliminary. Additional studies examining the effects of patient context on management decisions relative to CPGs among physicians from
a variety of disciplines, and from various communities are required to confirm the presented findings. Furthermore, a larger sample may provide opportunity to determine physician related factors associated with the proclivity to follow or deviate from CPGs when “social context” is present.

4.5 Conclusion

Clinical Practice Guidelines are developed to assist physicians in determining the best course of care for patients. While CPGs might provide opportunity to maximize outcomes at the population level, they are not designed to account for the unique needs of individual patients. The data presented here suggests that EM physicians are sensitive to both individual patient context and the best clinical evidence of benefit (as per CPGs) when determining the how care should be managed. Additional research is needed to determine the extent to which consideration of “social context” does result in better patient-important outcomes. If deemed important, the “social context” of patients should be considered when examining why physicians might not adhere to CPG recommendations. Furthermore, as many physicians in this study seemed to respond to context when determining how care would be managed, studies examining practice variation might benefit from a consideration of systematic differences in that context between populations when identifying the sources of variation.
<table>
<thead>
<tr>
<th>Case</th>
<th>Descriptor</th>
<th>CPG indicates</th>
<th>Social Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ankle 1</td>
<td>46 yr M trauma L ankle</td>
<td>No xray</td>
<td>Pursued by police</td>
</tr>
<tr>
<td>Ankle 2</td>
<td>21 yr M trauma R ankle</td>
<td>No xray</td>
<td>Pro hockey player</td>
</tr>
<tr>
<td>Ankle 3</td>
<td>36 yr F trauma R ankle</td>
<td>No xray</td>
<td>Civil suit</td>
</tr>
<tr>
<td>Head 1</td>
<td>62 yr M minor head inj</td>
<td>No CT</td>
<td>Lives alone</td>
</tr>
<tr>
<td>Head 2</td>
<td>47 yr F minor head inj</td>
<td>No CT</td>
<td>Chair of the hospital board</td>
</tr>
<tr>
<td>Head 3</td>
<td>93 yr M head injury</td>
<td>CT</td>
<td>Advanced alzheimer's</td>
</tr>
<tr>
<td>Strep 1</td>
<td>21 yr F sore throat</td>
<td>No Abx</td>
<td>Mother is a nurse who works in the ED</td>
</tr>
<tr>
<td>Strep 2</td>
<td>5 yr M sore throat</td>
<td>No Abx</td>
<td>Family members recently treated for strep, anxious parents</td>
</tr>
<tr>
<td>Strep 3</td>
<td>46 yr F sore throat</td>
<td>No Abx</td>
<td>Nurse practitioner, leaves for isolated community tomorrow</td>
</tr>
<tr>
<td>AF 1</td>
<td>54 y F palpitations</td>
<td>No warfarin</td>
<td>Vascular Surgeon</td>
</tr>
<tr>
<td>AF 2</td>
<td>64 y M palpitations</td>
<td>Warfarin</td>
<td>Alcoholic</td>
</tr>
<tr>
<td>AF 3</td>
<td>78 y F palpitations</td>
<td>Warfarin</td>
<td>Unreliable patient</td>
</tr>
</tbody>
</table>

Table 4-1: A brief summary of the survey cases and their associated social variables

<table>
<thead>
<tr>
<th></th>
<th>Experienced EM</th>
<th>Resident EM</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>Female</td>
<td>10</td>
<td>12</td>
<td>22</td>
</tr>
<tr>
<td>ED Census</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;25000 – 50000</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>&gt;50000-75000</td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>&gt;75000</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>mean (range)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years in Emergency Medicine</td>
<td>13 (3-33)</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Shifts/month</td>
<td>14.9 (6-36)</td>
<td>13.5 (4-16)</td>
<td>14 (4-36)</td>
</tr>
<tr>
<td>Patients/shift</td>
<td>31 (24-45)</td>
<td>12.7 (7-15)</td>
<td>21 (7-45)</td>
</tr>
<tr>
<td>Followed CPGs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV condition</td>
<td>3.2 (1-5)</td>
<td>4 (2-6)</td>
<td>3.6 (1-6)</td>
</tr>
<tr>
<td>No SV condition</td>
<td>4.6 (3-6)</td>
<td>4.5 (2-6)</td>
<td>4.5 (2-6)</td>
</tr>
<tr>
<td>SV/no SV &lt;1</td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 4-2: Participant demographics and characteristics of work environment. Rate of adherence to CPGs.
Figure 4-1. Rate of adherence to CPGs for both experienced and resident physicians, in both the SV and no SV conditions.
Chapter 5: Medical Practice Variations: What the Literature Tells Us (or Does Not) About What Are Warranted and Unwarranted Variations

Not all medical practice variations are necessarily bad for patient care. Discriminating between those that negatively affect patient care and those that exist for good reason is necessary if information from studies examining variation is to be useful in ensuring good care. However, it is difficult to account for desirable variation when designing studies (such as those presented in the previous two chapters) because previous studies are unclear as to when variation is warranted and when it is unwarranted. The study presented in this chapter describes a review of the literature regarding how “unwarranted variation” has been defined when examining medical practice variations. This study suggests that greater clarity is needed regarding this definition. A lack of a good definition can increase the opportunity for erroneous interpretation of study results when determining the implications of variation, and thus, is an important feature of variations research.

Medical Practice Variations: What the Literature Tells Us (or Does Not) About What Are Warranted and Unwarranted Variations

Mathew Mercuri MSc, PhD(C), Amiram Gafni PhD

McMaster University, Hamilton, Ontario, Canada

Corresponding Author:
Mathew Mercuri MSc PhD(C)
Centre for Health Economics and Policy Analysis, CRL Building, 282
McMaster University
1280 Main Street West
Hamilton, Ontario, Canada, L8S 4K1
Telephone: 905-525-9140 ext.22122
email: matmercuri@hotmail.com

Brief Title: What are Warranted/Unwarranted Practice Variations?

Word Count: 3243
5.1 Abstract

This paper examines the sources of practice variations and definitions of unwarranted variation, as derived from the literature. The literature suggests variables/factors related to patient health needs, physician “practice style”, and environmental constraints/opportunities as sources of practice variations. However, this list is likely to be incomplete due to significant unexplained variation in each study. Furthermore, it is unclear which factors are sources of unwarranted variation because the reviewed studies do not clearly discriminate between those variations that are unwarranted and those that are not. It is also unclear if context plays a role in determining if and when a factor is unwarranted. The literature contains few frameworks of what constitutes unwarranted variation. Among those offered, more information is needed regarding the scientific basis for including the selected factors, and how to operationalize the framework provided a particular one is chosen. A clear and consistent framework for unwarranted variation, and a clear indication how each component factor could be measured and integrated can help investigators determine which variables should be included in their studies, such that the sources of unwarranted variations may be identified. A better understanding of the role of patient preference as a potential source of practice variations is also required.

Key Words: Practice Variation, Unwarranted Variation, Practice Style
5.2 Introduction

Medical practice variations have been an important topic of interest among health services researchers ever since Glover’s systematic account of variations in the rate of tonsillectomy procedures between geographic areas in the United Kingdom.⁵ At the time, it was unclear as to the reasons for such observed variations. Decades later, led most notably by Wennberg, investigators would demonstrate practice variations in a wide variety of medical procedures and treatment strategies.²⁻⁶ These investigators attributed this variation to a multitude of sources, including the individual physician managing the care and the environment in which the care is delivered.

To some, the very existence of medical practice variations implies that some individuals are not receiving the best care or that healthcare resources are being used inappropriately. In that case, the primary goals of practice variations research would be to both reveal the existence of variations and provide information on which factors to target with interventions, such that these variations may be minimized (and care optimized). However, some variations may exist for good reasons. For example, variations occurring because of differences in patient needs might be considered warranted. Few (if any) would argue that such “warranted” variations should be minimized. Instead, only those variations that have a negative impact on care should be eliminated. Therefore, to meet the goal of eliminating unwarranted variations in care one requires knowledge of all the factors that cause them to exist and then define/determine those that are warranted and those that
not. Furthermore, we must determine if these factors are universal or contextually based. That is, it might be that in a certain context a given factor is a cause for unwarranted variations and in another context it leads to warranted variations.

The paper presented here examines the literature on medical practice variations and will proceed in two sections. First, we highlight the literature on medical practice variations to date. This will include an identification of the professed sources (both hypothetical and empirically demonstrated) of previously observed practice variations. In the second part of this paper we will examine the concept of unwarranted variation. This will include a review and appraisal of various frameworks of unwarranted variation proposed in the literature, and suggestions for future research.

5.3 Literature Search

A literature search was performed to identify articles regarding a framework of the term “unwarranted variation”. This was done using the PubMed, and Google Scholar electronic databases. "Unwarranted variation", “unwarranted practice variation”, and “unwarranted medical practice variation” were used as keywords in the literature search. This search yielded 511 unique articles. Article titles and abstracts were read to assess their relevance for this review. Only those articles published in English and which unwarranted variation was used in reference to medical practice variation were deemed relevant. The reference sections of relevant articles were then reviewed to identify additional articles that were not found
during the database search. In total, 46 articles using the term “unwarranted” in reference to medical practice variations were identified. The term “unwarranted” was chosen because it seemed to be the term most commonly used to distinguish “appropriate” variations from “undesirable” variations among those articles identified in the general search on the causes of medical practice variations. We did not formally search the databases for articles using somewhat synonymous terms, such as “unacceptable”, “inappropriate”, or “unjustified” because these terms rarely appeared in the reviewed literature on the sources of medical practice variations. Furthermore, our limited attempts to identify relevant articles using these terms yielded few results. Thus, we singularly focused our attention on the more commonly used “unwarranted” as to avoid confusion when examining the conceptual meaning in reference to variation.

5.4 Sources of Practice Variation

5.4.1 Types of Practice Variation

There are many types of practice variation. Early research focused on variations in the rate of use of common surgical procedures between geographic areas. These were collectively known as “small area variations”. Small area variations studies implied that the care an individual receives is contingent on where and by whom the care is provided. It also assumes, implicitly, that for a given indication, in a given population there is a “right rate of use” of a procedure or an intervention. Thus variation in use might indicate under or over use of the
procedure/intervention in certain areas if the right (or optimal) rate of use was known. More recently, researchers have noted practice variations at different levels of aggregation. This includes variations between hospitals or groups of physicians, and between individual physicians of similar expertise, in addition to variations between small geographic regions (i.e. small area variations). In addition to the differing types of practice variations, it is also unclear as to how much of a difference between units is required for it to be considered a practice variation. The literature does not offer a definition as to what are the criteria to determine when a difference can be labeled as practice variation (statistical significance, minimum number varying from the “pack”, or other). As it is not the purpose of this paper to define practice variation, we will accept the authors “assertion” regarding the existence of variation and examine the issue more generally. We will consider practice variations at all levels of aggregation and significance (both statistical and clinically relevant).

5.4.2. Uncovering the Sources of Practice Variations: Early Research and the Role of the Physician

The modern era of medical practice variations research began with the seminal work by Wennberg and Gittelshon, in which the authors demonstrated the existence of geographical variations in the rate of use of different surgical procedures. Furthermore, the authors recognized that these “small area variations” persisted despite adjusting their rate estimates for age. Numerous
studies would later confirm that even after case-mix adjustments, including prevalence of disease, patient demographics, and social-economic factors, such as income, medical practice variations are still observed for many medical procedures and therapies.⁴,⁷-¹⁴

Realizing that patient case-mix did not fully explain practice variations, researchers began to consider the role of the physician as an additional source of variation. Wennberg was the first to introduce the idea that medical practice variations (specifically “small area variations”) may in part be caused by the physicians themselves.¹⁵ He termed this factor “practice style”.¹⁵ Although “practice style” is not clearly defined in the literature, researchers have offered numerous hypotheses regarding its “causes”.¹⁶-¹⁸ Regardless of how one defines “practice style”, there is evidence to suggest that physician-related factors do influence treatment decisions. These include fear of malpractice, how the physician interprets the available evidence/clinical information, training environment, clinical knowledge, threshold for action, and tolerance of specific risks.¹⁹-²⁸ Because the magnitude and relative importance of these factors likely differs among physicians, it is reasonable for one to expect heterogeneity in treatment decision between physicians given the same case(s).
5.4.3 Criticism of the “Practice Style” Explanation and the Role of Contextual Constraints

Overall, the evidence in favour of “practice style” as the primary cause of observed medical practice variations is somewhat mixed. There are studies that both support (e.g. 29-31) and refute (e.g. 32-36) this claim. The main criticism may be that previous studies are inadequate in their case-mix adjustments, as most did not measure or account for potentially important patient characteristics, such as disease severity. Furthermore, as Stano points out, a considerable amount of the variation remains unexplained in studies exploring medical practice variations (specifically those exploring small area variation, but the argument applies to other levels of aggregation and current studies as well), and thus, it is premature to attribute this unexplained variation to “practice style”.16,37

Acknowledging these limitations, some researchers sought out another way to conceptualize the problem of practice variations. The role of “practice style” would be de-emphasized in favour of contextual and environmental constraints acting on the physician. This approach would generate such conceptual frameworks as the “resource demand model”, and the “constraint-centered model”.38-40. While innate physician characteristics would still account for some of the observed variation, upon adequate measurement of and adjustment for patient characteristics and organizational/environmental constraints, little variation between providers or geographical regions should exist. Some studies do implicate supply factors and environmental constraints as the primary source of practice variations (e.g. 35,41-43).
However, as with studies attributing the observed variation to “practice style”, a significant amount of the variation remained unexplained.

5.5 Are All Variations Bad?

Over the past decade, researchers began to recognize that 1) part of the observed variations may exist for good reasons, and 2) it is only those variations that may have a negative effect on patient care that should be minimized.44 Traditionally, studies did not distinguish between those variations that exist for good reason and those that do not. This may be problematic as some stakeholders may use information extracted from these studies as a framework for interventions designed to minimize variation, hence implicitly assuming that they are not warranted. However, in order to determine if a variation (and its sources) is acceptable or not we first need a good understanding of what constitutes an “unwarranted variation”. A good framework to determine “unwarranted variation” should provide: 1) a scientific basis (e.g. theory, conceptual framework) to justify the inclusion or non-inclusion of each factor, and provide reason for when it is and is not applicable, 2) a clear definition and explanation of each component or factor that is suggested as a cause for unwarranted use, and 3) explain how it is to be operationalized – including how each factor is to be measured, and, if there is more than one factor, how to integrate them.

Below we provide a review of how the term “unwarranted” has been used in the literature in reference to medical practice variations. The results of our review
are organized in the following manner. We begin by presenting articles where the authors provide a framework of the term “unwarranted variation”. This includes a comparison of the proposed frameworks, along with a brief appraisal based on the criteria provided above. Next, we present articles where even though “unwarranted variation” is not explicitly defined a framework is implied by the authors. These “implied frameworks” will be compared with those that were more explicitly stated. Thirty-two of the forty-six articles identified do not offer a framework of unwarranted variation. However, in fourteen such articles the author(s) reference one of the frameworks we will be reviewing.\(^{45-58}\) The remaining 18 articles do not offer a framework or provide a reference for the use of the term “unwarranted variation”. Only 6 of these articles\(^{59-64}\) report on studies to identify medical practice variations, whereas the other 12 were editorials, commentaries, review papers, or research studies examining healthcare without specific attention to identifying practice variations.\(^{44,65-75}\) As these 32 articles do not offer criteria on what is or is not unwarranted variation, they will not be discussed further in this review.

5.5.1 Suggested Frameworks of Unwarranted Variation in the Literature.

Wennberg seems to be the first to provide a framework for the term “unwarranted variation” in reference to medical practice variations.\(^{76}\) The variations he describes are “unwarranted because they cannot be explained by type or severity of illness or by patient preferences”.\(^{76}\) Wennberg further divides unwarranted variation into 3 categories (i.e. effective care, preference sensitive
care, and supply sensitive care), believing that their causes and subsequent remedies differ according to category. He would reiterate this perspective in two more articles on practice variations, in one case stating a more patient focused framework of unwarranted variation: “care that is not consistent with a patient’s preference or related to a patient’s underlying illness”. In an interview, Wennberg offered a similar framework, where unwarranted variations are those that “cannot be explained on the basis of illness, patients’ preferences, or dictates of scientific medicine”. While Wennberg’s framework appears to contain several aspects that he feels affect care, what is unclear is why these specific factors were chosen and not others, whether those chosen are mutually exclusive, and if not, how to integrate them when attempting to identify unwarranted variations in clinical practice.

Likewise, Wennberg does not offer specific instruction on how to incorporate his (or any) framework of unwarranted variation into studies to identify their existence/sources. How to do so may not be obvious for each of the factors he has presented. For example, it is not clear if patient preferences can be truly measured using current methods, and provided that they indeed do, such methods are resource intensive - making them difficult to incorporate into practice variations studies looking at large populations.

Three more frameworks were identified in the literature. Sepucha et al. suggested, “there are multiple sources of unwarranted variation including inequitable access to resources, poor communication, and role confusion. However, perhaps the most significant source is the misinterpretation or misapplication of the
relevant clinical evidence.”. This framework seems consistent with the part of Wennberg’s framework dealing with “effective care”, and perhaps even that covering “supply sensitive care”. Although not explicitly stated as part of their framework, Sepucha et al. further suggest that “the degree to which care should vary based on patient preferences is equally important”, implying variations due to differing preferences in care should be considered warranted. Goodman suggests “an expanded definition of unwarranted variation is that it is the variation in medical resources, utilization, and outcome that is due to differences in health system performance”, and that “‘Unwarranted’ refers to the portion of the variation that is explained not by population difference but by the quality, appropriateness, and efficiency of health care.”. This framework seems consistent with those mentioned above (in fact, Wennberg’s three categories of variation are described by Goodman). However, it is not entirely clear what is meant by “quality” or “efficiency”, or how these should be measured. The last framework of unwarranted variation we identified is that by Bojakowski, who claimed that “unwarranted variations in healthcare services are variations that cannot be explained by public needs or medical needs”. Unfortunately, this framework may not be useful for decision-makers and researchers, as need is difficult to operationalize. Furthermore, as Bojakowski points out, “whether a variation is deemed to be acceptable, however, will ultimately be a matter of judgment”. This only adds to the confusion. Thus, what Bojakowski believes represents unwarranted variation,
and whether this framework is consistent with those proposed above is difficult to determine.

Overall, there appears to be some agreement among the proposed frameworks regarding what constitutes “unwarranted variation” in medical practice. Variations in the uptake of therapies accepted to be most effective is generally considered unwarranted. Likewise, variations explained by patient preferences should be taken into account prior to classifying observed variation as unwarranted. The authors, however, do not offer a clear scientific basis for why these specific factors should be included and not others. Furthermore, the authors do not offer guidance in how to operationalize their proposed frameworks when conducting practice variations research. Clearly, there may be differences in opinion on what constitutes effective care or patient preference and how to measure these effectively.

5.5.2 Articles Implying a Framework

Seven articles in which the authors seem to imply a framework of “unwarranted variation” were identified.\textsuperscript{83-89} The first example comes from a document describing the role of the National Institute for Health and Clinical Excellence (NICE) in improving healthcare in the United Kingdom. One of the core functions of NICE is to “reduce unwarranted variation in practice across the U.K., through the development and dissemination of best practice evidence-based standards”.\textsuperscript{83} Although a framework of unwarranted variation is not explicitly
defined in this statement, the method(s) through which NICE intends to reduce this variation alludes to its meaning. It would appear that unwarranted variations are those differences in care that exist despite compelling evidence or agreement among providers as to the (evidence based) best course of treatment. Likewise, most of the identified implied frameworks appear to suggest that variations caused by deviations from evidence-based practice guidelines or failure to provide therapies supported by clinical trials are unwarranted.\textsuperscript{84,86-88} One author implies that variations caused by non-clinical factors, such as type of insurance or treatment location, are unwarranted.\textsuperscript{85} The articles presented above do not address the role of patient preference with respect to unwarranted variation.

5.6 Discussion

The usefulness of information derived from practice variations studies is somewhat predicated on the ability of those studies to determine both the existence and the causes of unwarranted variations. Failure to discriminate between variations that are warranted and those that are not could result in more harm than good in the event that strategies/interventions are designed to eliminate all observed variation. That is, such interventions risk eliminating variation that existed for good reason. However, discriminating between warranted and unwarranted variation appears to be problematic. Although discussion has begun over that past decade regarding a framework of what constitutes “unwarranted variation” there is some disagreement as to which factors should be included. In the
event that we adopt one of the more comprehensive frameworks (such as that by Wennberg), it is still unclear what is the scientific basis for the factors identified, and how to integrate the different factors when attempting to use the framework in the context of an empirical study that seeks to identify the magnitude and sources of unwarranted variation.

Regional or provider differences in the rate of use of “effective” treatments supported by trial evidence appear to be the main focus of previous studies in medical practice variations. This is the case for studies published both before and after the arrival of frameworks of unwarranted variation in medical practice. Patient preference has received little attention in this area of research. To our knowledge, patient preference is rarely measured when examining practice variations. The single study that we did identify examined the role of patient “care-seeking” preference in determining regional variations in healthcare use.\(^{90}\) Therefore, it is difficult to determine if previously observed practice variations are indeed “unwarranted”, at least in the manner proposed by the frameworks presented earlier. This fact may speak to either the difficulty in operationalizing these frameworks when conducting an empirical study attempting to measure medical practice variations, or to disagreements among the research community regarding the role of patient preferences in explaining such variations. Patient preference is a potentially important issue, as some who advocate evidence-based medicine now recognize patient preferences as a justifiable source of variation.\(^{91}\)
Which framework of unwarranted variation investigators should ultimately adopt from those suggested here or from new ones that need to be developed remains unclear. Regardless of which is ultimately chosen, the framework must 1) be clear which items represent sources of warranted (or unwarranted) use, 2) justify their inclusion (i.e. provide a scientific basis for their inclusion), and 3) define each component factor so that they can be measured. However, the integration of these items (e.g., the relative importance) may be context specific. Therefore, in addition to the above criteria the framework should provide guidance as to how to take context into account when integrating the different factors. Currently, the information required to sufficiently follow this process cannot be obtained from the published literature.

5.7 Conclusion

The arguments presented in this paper evince the need for a clear and consistent framework of what constitutes unwarranted variation. This may be difficult because the meaning of unwarranted variation is seemingly a moving target. This paper also points to the need for better information regarding which factors should be included in studies once a particular framework is chosen. Improved understanding in these two areas may allow researchers to learn when unwarranted variations exist, and what are its causes. This information can assist healthcare providers in optimizing strategies designed to minimize unwarranted variation.
5.8 References


Chapter 6: Using Small Area Variations to Inform Healthcare Service Planning:

What Do We “Need” to Know?

Earlier chapters focused on determining the sources/mechanisms by which variation arises. While it is necessary to understand why variation exists, and when it is undesirable, it is also important that what is learned from these studies can provide those entrusted with planning healthcare services and allocating resources with valuable information to assist them in doing so effectively (i.e. maximize goals as determined by the policy under which they operate). Thus, from a healthcare services perspective it is important that the focus of variations studies is aligned with the goals of healthcare managers. The study presented in this chapter suggests that these goals are not necessarily aligned, and thus, their usefulness to healthcare managers may be inadequate. This study argues that incorporating population needs into variation study designs constitutes an important and necessary methodological advancement.

Using Small Area Variations to Inform Healthcare Service Planning: What Do We “Need” to Know?

Mathew Mercuri PhD(C), Stephen Birch PhD, Amiram Gafni PhD

McMaster University, Hamilton Canada

Correspondence to:

Mathew Mercuri

Centre for Health Economics and Policy Analysis
McMaster University
1280 Main Street West
Hamilton Ontario Canada
L8S 4K1

Email: matmercuri@hotmail.com

Word Count: 4095

Short Title: Variations and Healthcare Policy

Keywords: Small Area Variation, Resource Allocation, Use, Need, Healthcare Policy
6.1 Abstract

Rationale, aims and objectives: Allocating resources on the basis of population need is a healthcare policy goal in many countries. Thus, resources must be allocated in accordance with need if stakeholders are to achieve policy goals. Small area methods have been presented as a means for revealing important information that can assist stakeholders in meeting policy goals. The purpose of this review is to examine the extent to which small area methods provide information relevant to meeting the goals of a needs-based healthcare policy.

Methods: We present a conceptual framework explaining the terms ‘demand’, ‘need’, ‘use’ and ‘supply’, as commonly used in the literature. We critically review the literature on small area methods through the lens of this framework.

Results: ‘Use’ cannot be used as a proxy or surrogate of ‘need’. Thus, if the goal of health care policy is to provide equal access for equal need, then traditional small area methods are inadequate because they measure small area variations in use of services in different populations, independent of the levels of need in those populations.

Conclusions: Small area methods can be modified by incorporating direct measures of relative population need from population health surveys or by adjusting population size for levels of health risks in populations such as the prevalence of smoking and low birth weight. This might improve what can be learned from studies employing small area methods if they are to inform needs-based healthcare policies.
6.2 Introduction

Whatever resources are available for healthcare, this level remains insufficient to meet all the potential demands for services that managers face. Thus, it is desirable that available resources be allocated in a manner that optimally meets the goals of the healthcare policy. A goal that is often cited is to maximize the health benefits accruing to the community. Over the past few decades, variations in the use of healthcare services have been observed between geographic areas. These so-called “small area variations” (SAV) raise concerns that resources are not distributed equitably or in a manner that does not maximize health benefits for a given population. To some, the existence of variation implies its own solution, namely, reducing variation, and this notion has played an important role in healthcare reform over the past few decades, most notably in the United States [1]. Wennberg, has championed the cause for reducing geographic variations, arguing that the elimination of variation will improve efficiency, and thus, increase health outcomes [2]. Wennberg recently argued for routine reporting of variation, presumably based on small area methods, to inform policy decisions and healthcare providers [2]. Current interest in the policy relevance of geographic variations is also evident in the recent call for papers by Health Policy for a special issue celebrating 40 years of “small-area variation” [3].

While few would argue against eliminating variations that reflect suboptimal resource use as defined by policy goals, the utility of applying traditional small area methods to inform policy makers is questionable, especially when policy goals are
focused on maximizing health gains by aligning access to care with needs for care.

In particular, SAV studies measure use of services in different populations independent of the levels of need in those populations. But use and need are different concepts, so use is not a valid surrogate measure or proxy for need. In other words, a use-based allocation of resources is unlikely to be the same (or even similar) to a need-based allocation of the same resources.

The necessity to obtain information on population needs rather than use for maximizing health gains to the population has long been recognized. For example, shortly after Wennberg and Gittelsohn's [4] landmark paper on SAV, the United Kingdom’s Department of Health and Social Security issued their Report of the Resource Allocation Working Party (The RAWP Report), which outlined a methodology to base resource allocation on relative levels of population needs. How to account for population needs in resource allocation formulae became a topic of interest in the health policy literature at the same time as studies of SAV flourished (e.g. [5]). However, while these two bodies of literature grew concurrently, SAV methods developed somewhat independently. As some authors advocate that SAV studies should guide resource allocation, it is important to understand what they measure, and how this information is relevant with respect to healthcare policy goals. Furthermore, it is important to critically appraise SAV methods, as few have questioned the validity of their conclusions or usefulness in guiding resource allocation [6].
In this paper, SAV methods are examined for their use in informing healthcare policy. First, we present a conceptual framework explaining the terms “demand”, “need”, “use”, and “supply” as commonly used in the literature. Next, we critically appraise the role of SAV studies in informing healthcare policy. Finally, we offer suggestions for modifying the SAV approach to provide information necessary to inform healthcare policy.

### 6.3 Conceptual Framework

#### 6.3.1 Demand, Need, Use, and Supply

In the context of healthcare, demand “represents the request of the patient (or the physician or the patient’s family) for medical services” [7, p.71]. The extent to which an individual will request (and potentially use) healthcare services is contingent on his or her perceived needs, predisposing characteristics (e.g. demographics, social structure, health beliefs), and resources [8]. By “need”, we mean the morbidity (illness) burden of an individual or defined population that is medically modifiable [7]. The extent to which the responsible physician (or other healthcare provider) will request healthcare services on behalf of an individual is subject to the individual consulting the provider and the provider’s evaluation of that individual’s medical needs, perception of that individual’s preferences, and the availability/supply of appropriate healthcare resources. By supply we mean both human (e.g. physicians, nurses, etc.) and non-human (e.g., buildings, equipment and consumables including medicines) resources used for providing a medical/surgical
consult, procedure or intervention. The “use” of healthcare services is the interaction between demand and supply as illustrated in Figure 6-1. In standard economic theory of competitive markets, supply and demand are independent. However, in healthcare markets an individual uses the medical expert to advise on which healthcare services “should” be demanded in order to achieve the individual’s desired health change. Thus, the provider’s influence can be used to shift the demand schedule in order to manage his/her workload and income expectations (Figure 6-1). In this way the supplier can induce a demand for healthcare (as distinct from an individual’s demand for health change). As a result use may not reflect only need (e.g. supplier induced demand), nor will need always be reflected in use (e.g. barriers to access to care).

The distinction between use and need has received attention in the medical literature [e.g. 9-11]. Most notable was Tudor-Hart’s [9] discourse on the “inverse-care law”, whereby the availability (and by extension, use) of healthcare tends to vary inversely with population needs. This may be the case even in systems with universal coverage and first dollar public funding of healthcare. For example, the introduction of the UK National Health Service in 1948 essentially nationalized payments for healthcare with care being delivered through the existing configuration and distribution of the healthcare infrastructure, which reflected pre-NHS market forces (i.e. willingness and ability to pay for care). So although individuals no longer faced healthcare prices as a barrier to accessing care, they still had to find a source of care. Because poorer communities had a relatively lower
level of supply than more prosperous communities, but relatively higher needs for care, social inequalities in access to care when needed remained. Thus, a healthcare infrastructure, inherited from a pre-government-funded healthcare era is perpetuated despite government insurance because the care is more easily available (and hence use relative to need is greater) among those populations who traditionally could afford it rather than those populations who need it most.

Likewise, others have argued that healthcare resource (e.g., hospital) use is a poor proxy for need (e.g., [10, 11]) because of pre-existing variations in supply.

Empirical evidence exists to support this distinction between use and need in real practice. For example, when examining the determinants of family physician use in Canada, Birch and colleagues [12] showed that the association between levels of use and levels of need is significantly related to a range of other variables unrelated to need. Meltzer and Hochstim [13] report a substantial gap in the health status of survey respondents and what appears in their medical records, especially among chronic diseases. This could be because 1) some patients are not using (due to barriers or preference) healthcare services or not perceiving the need for such in the case of some conditions, despite poor health status, or 2) the individual presents for care, but after evaluation the physician does not think it warrants further service use.
6.3.2 Healthcare Policy Goals

No matter what level of resources are made available for healthcare, supply will be insufficient to meet all the potential demands for care. Healthcare managers must set priorities for supplying care (i.e., allocating health care resources) in accordance with the healthcare policy goals of their community or government. In many countries, the goal of healthcare policy is to maximize the impact of public healthcare resources on population health need, which requires that those resources be allocated according to medical necessity [14,15]. Explicit needs based approaches to resource allocation have been adopted in several European countries (e.g. Austria, U.K, Finland, Sweden, Germany) [16-20]. Alternatively, healthcare resources might be allocated according to demands for care (as influenced by supplier recommendations). For example, in Canada, among healthcare services not covered in the Federal Governments Medicare legislation, such as dental care, chiropractic care, and physiotherapy, individuals’ ability and willingness to pay for care (either out of pocket or through private insurance), together with what the provider recommends will determine how resources are allocated.

6.4 Small Area Variations: What they measure, and how they might *(mis)inform* Healthcare Policy

Beginning with Wennberg and Gittelsohn’s [4] study, use of healthcare was the dependent variable in SAV studies. Indeed, Wennberg later confirmed that “most of the subsequent work in small area methods has concentrated on the
utilization of care” [21, p.YS76]. Researchers have reported variations using one or a combination of variables, such as hospital admission rate, surgical procedures performed, expenditures, and a variety of available resource metrics (e.g. beds, physicians, nurses, etc.). Traditionally, SAV are reported as a population based rate (e.g., hospital admissions per 100,000 population) in the community or geographical region, a proportion of patients receiving some therapy (from a defined set of patients accessing services) (e.g. [22-24]), or mean levels of use (e.g. length of stay, cost per case, etc.) (e.g. [25]). Often these rates are adjusted for population characteristics such as age and sex (or in rarer cases, disease severity, co-morbidities, etc.) (e.g. [4,26,27]). In each case the dependent (outcome) variable(s), such as those summarized here, describe use of services and/or resources.

Wennberg argued that information on practice variations should be routinely reported, as it “is important for examining the relations between policy decisions and clinical decisions and raises important questions concerning the efficiency and effectiveness of healthcare” [2, p.687]. The intended role of SAV studies in informing policy goes beyond simple reporting. Wennberg suggests that “small area analysis provides an important tool for the reform of health care markets because it poses and helps answer the ‘which rate is right?’ question” and “it provides information relevant to the control of capacity” [21, p.YS75]. Wennberg also argued that evaluative clinical sciences (including SAV methods) play an important role in the healthcare agenda as “they expose unwarranted variations in care and can be used to remedy them” [28, p. 961].
Others have made more explicit recommendations on how SAV studies can inform healthcare policy. Recognizing the role of reducing/containing escalating healthcare costs as an important aspect of the policy agenda, Paul-Shaheen et al. suggest “small area analysis offers health policymakers an opportunity to target cost-containment efforts to reduce selectively the use of unnecessary or ineffective services while promoting the delivery of those services making a positive contribution to health” [29, p.771]. Specific recommendations for policy include promoting a research agenda to identify the magnitude of inappropriate care where variations exist, advancing strategies to reduce professional uncertainty and improve uniform professional decision making rather than concentrating deviant behaviour, and promoting the role of education and provider feedback [29].

The importance that SAV studies have on informing policy is dependent on how they are interpreted (e.g. which factors they attribute as the cause). There are 3 general sources of variation: patient level factors, provider level factors (including individual physician “practice style”), and practice context/environment [30]. Although these factors are known generally, many studies suffer from much of the variance remaining unexplained, and thus, the residual variation tends to be attributed to one factor or another. Giving priority to one of these factors, for example physician “practice style” [31-33], will result in different strategies to minimize variation (e.g. academic detailing, practice guidelines, report cards), than it would if one were to give priority to a different factor, such as supply or environment [34] (e.g. promote changes in resource allocation to minimize
variation). Specific factors aside, if one were to simply interpret SAV as an indication of inequitable care (i.e. different treatment for same need), strategies required to align service provision with a healthcare policy based on needs would necessitate focus on elimination of all SAV (possibly after adjustment for population size and demographic mix), generally speaking.

If we consider a healthcare policy based on needs (this is an important component of many policies, as mentioned above), then information from studies employing traditional SAV methods (based on use) is inadequate in informing healthcare policy because use is a poor surrogate measure of need. For example, use overlooks those who need care but do not use it, nor can we adequately determine which individuals among those who use care do not need it. The implications of the failure to consider the needs of the population when observing/interpreting SAV become more evident when we look at how needs are related to what small area methods measure. Consider that small area methods measure the quantity of services used in a defined population (Q) per unit of that population (P). Incorporating needs (N) yields equation 1 (below). Suppose the ratio Q/P is held constant (i.e. two populations have the same rate of use). In this case, the population having higher relative needs (i.e. higher ratio of N/P) would thus have a lower rate of use per unit need (lower ratio of Q/N). Such a scenario would be counter to a needs based healthcare policy, where resources are to be allocated on the basis of equal use for equal need (i.e. equality between populations in the ratio of Q/N). Furthermore, attempts to equalize the ratio of Q/P in two
populations that differ in this regard (i.e. eliminating a small area variation) would risk creating precisely the scenario we describe above, in the event that we do not know the needs of the populations in question.

\[
\frac{O}{P} = \frac{O}{N} \times \frac{N}{P}
\]  

[equation 1]

Proponents of small area methods acknowledge to some extent that geographical differences in need exist. However, in most cases the adjustments rely on demographic differences (e.g. age and sex), which do not reflect needs (or even health status – see [35]). Two communities with different age profiles do not imply that the older community has greater health needs. The “younger” community may have a higher level of age-specific needs, such that there is substantial premature death (thus lowering the average age of the population). The higher level of (age specific) need might warrant higher (age specific) use of care but adjustment for age would simply exaggerate the problem as the ‘younger’ community’s resource per capita resource share is reduced. Any observed higher level of use by the younger community would be interpreted as “overuse”, under the SAV approach. Attempts to eliminate the observed variation through demographic adjustments to resource allocations would result in poorer outcomes overall, and would be counter to a healthcare policy aimed at maximizing health gain, because those who need it most were not old enough on average to “require” it. In other words, SAV studies do not
provide adequate information to determine which is the “right” rate of use on the basis of needs. Furthermore, one must be cautious in assuming there is a universal “right” or “correct” population rate of use, rather, there is a socially optimal rate that is context specific. What the example above illustrates is that if the needs between the two communities differ, then the “right” population rate of use in each community will also differ.

If policy were indeed concerned with provision of service on the basis of equal access/use for equal need, appropriate planning would eliminate only those variations that exist without a justification on the basis of need. Although recognized by proponents of small area methods [36], past studies do not distinguish between variations that are “warranted” and those that are “unwarranted” [30]. Without this distinction, one might assume, perhaps incorrectly, that all observed and unexplained variations are “unwarranted”. Under a needs-based healthcare policy, an “unwarranted” variation might be one whereby access/treatment is not equal for same need, but to know this researchers must measure need.

Focusing attention on use in populations without considering needs affects the interpretation of variations data for informing policy and service provision in other ways. Small area variations in hospital admissions and surgical rates might reflect a substitution of institutional care for community services. For example, patients in one region may exhibit a longer than average hospital stay or may be admitted more often if other social services, such as long term care homes,
outpatient rehabilitation centres, home care services, or hospice, are less available in the region compared to other regions. Such factors are not readily measured using traditional SAV methods. It could be that different communities are meeting the policy goals in different ways, and yet data acquired using SAV methods could imply that some are not.

If policy goals are based on something other than needs, such as demands, small area methods may be no more appropriate because SAV methods do not measure variations in demand, but rather, variations in use (i.e., the intersection of demand with supply use) (see Figure 6-1). Hence, SAV would not inform policy makers pursuing non needs-based goals either. In the event that use is largely a reflection of supply, and the population distribution of supply is largely unregulated, then SAV are likely to represent differences in the subjective value of providers working in those areas and their preferred workload.

6.5 Modifying Small Area Methods to Inform Policy

Tanenbaum argues that “variation is not a palpable problem; it is a statistic from which inferences are drawn, and the strength of that statistic – the information it does or does not provide – would seem to determine the potential benefits of reducing it” [1, p.11]. Despite the policy interest in variation, and seemingly general acceptance that it should be reduced [1], variations in medical practice continue to abound. It may be that the variation “statistic” is too general, and that researchers have yet to provide practical metrics for determining the relative influence of its
sources (i.e. patient, physician, resources), such that policy makers can translate information from variations studies into equitable and efficient use of resources. On the other hand, it could be that researchers are measuring the wrong thing altogether - the “statistic” itself is providing the incorrect information for our policy needs (i.e. variation in use rather than need).

Research examining SAV developed from observations of differing rates of use between populations, based on analysis of readily available information. The methods evolved as researchers recognized that populations differed in needs, incorporating available demographic information as a proxy for relative need. While this improved what could be learned from SAV methods, demographic information can be a poor indicator of need (as we describe above). Again, this is especially problematic where one desires to use data derived from SAV research to inform a needs-based policy.

Further improvements to traditional SAV methods might be achieved by incorporating community level indicators of population health status. For example, low birth weight usually involves greater service intensity at birth and in the perinatal period and has also been shown to impact visual, cognitive, and motor skill development [37]. The result is both poorer education outcomes, and higher need for healthcare services among those with low birth weight when compared to those with a normal birth weight [38-40]. Thus, low birth weight might indicate, in part, the health status of the population, where populations with a higher proportion of low birth weight individuals might be expected to have higher healthcare needs (all
other factors being equal). As the majority of births in Western Europe and North America (regions where SAV have received a great deal of study, and where needs is part of a healthcare policy) are assisted by a healthcare professional, it is likely that information regarding the incidence of low birth weight might be obtained from birthing records, clinic databases, or insurance records. One could adjust the population estimates of healthcare services use according the prevalence of low birth weight. Standardized morbidity or mortality ratios might also be used as an indicator of population health status [41,42]. While such measures might not give an accurate estimation of the absolute level of population needs, including them in small area analyses might offer some improvement in our understanding of relative differences in health status and risks to health between populations compared to analyses that consider demographics (most notably age and sex) alone.

Where indicators of population health status are not available or comprehensive, one could incorporate information on behavioural and social determinants of health. For example, one could adjust allocations for smoking status, socioeconomic status, education outcomes, such as the rates of post-secondary education, etc. These measures are often found in census data, population surveys, or other sources outside hospital records [43]. Small area methods could also incorporate information on non-hospital services that may affect how hospital services are used. This could include accounting for outpatient clinics, and hospice, and for other, non-medical, community resources.
Ideally, one could record the health status of each individual through an additional question on the national census, or population-based survey. However, census studies are infrequent, and population-based surveys have limited capacity for small area analysis because of sample size constraints. Where such information are available, Le Grand [44] offers a method to incorporate this information, whereby healthcare use (or cost) per needs ratios are calculated and compared between populations, adjusted for age and sex. Although this study was interested in inequities on the basis of income groups, one could extend the analysis to compare geographic regions. The method assumes between-group homogeneity in the distribution of disease as a proportion of the total morbidity within each group. If this is not the case, additional adjustment for disease distribution will be required. Collins and Klein [45] offer an alternative method for calculating use per unit need, again incorporating survey data regarding health status of the populations under observation [46]. Such methods have likely not been pursued due to data limitations and the necessary cost to acquire the relevant information.

Even if we can account for needs, observed SAV will at best only offer direction in how to allocate a defined set of healthcare resources equitably between regions/populations (i.e. equity in rate of use per unit need). What will not be apparent is which context specific rate is the correct one. Which is considered correct for each population may depend both on the goals of health policy, and on local characteristics of each population [47,48]. If the policy is to align resources with needs, then the correct rate of use in each context is the one where the region’s
The utilization of services is commensurate with its needs. If the needs differ between populations, the correct rate of use for each population will also differ. Thus, the notion of a universal “right” or correct rate of use is a fallacy.

6.6 Conclusions

The impact of SAV studies on our understanding of health services is undeniable. However, what they reveal about how we should allocate resources is limited if we are concerned about maximizing the health gain produced from available resources. Small area variation studies reveal differences in resource utilization between what appear to be similar geographic areas. Without information on the healthcare needs of the communities observed, using SAV studies to inform a healthcare policy based on need must assume that 1) all that need care in the population access it, 2) those that do not need care do not use it, 3) all those accessing healthcare get the best treatment to meet their needs, and/or 4) the different geographic areas defined within a SAV studies do not differ with respect to these assumptions. Unfortunately, these assumptions may be false. The current approach to medical research (i.e. evidence-base medicine paradigm) recognizes the possibility of confounding factors in any observational design (thus, the basis for performing randomized-controlled trials). Without an understanding of and adjustment for what might affect use of services beyond health status alone, we cannot say if any of the observed geographic areas is providing the level of utilization that will reflect the true needs of its population.
The literatures on needs based allocation and SAV have developed concurrently over the past 40 years. However, the integration of population needs into small area methods has been minimal. Small area methods should reduce the emphasis on data collected among those who access hospitals, and incorporate information from the general population each “small area” serves. This might reduce the residual confounding effect of unknown factors on our observations of SAV, ultimately giving a better estimate of their existence. In this way, the focus of SAV studies can move from utilization of services to needs for services, which might give stakeholders the information they need to improve resource allocation/distribution and service provision to meet policy goals based on needs.
Figure 6-1:

Use (U) is the intersection of the Supply (S) and Demand (D) schedules. The supplier can induce a change in the Demand schedule (D → D'). This will cause a subsequent change in the location of the intersection of the Supply and Demand schedules, such that U → U'.
6.7 References


Chapter 7: Discussion and Conclusions

7.1 Summary of Important Findings

Medical practice variations are ubiquitous. The recognition of variation that cannot be fully explained by differences in patient health status has raised questions about the appropriateness of both physicians’ practice and how resources are allocated. Where variations exist, physicians are viewed as lacking knowledge of how to best treat their patients, or insensitive to patient needs for the purpose of meeting some personal goal. In addition, resource allocation is viewed as inefficient. The impact of this perspective is undeniable. Where variation is seen as due to knowledge gaps, clinical trials and practice guidelines are advocated. Certainly the rise in evidence-based medicine (EBM) coincided with the increase in studies reporting variation, and in many ways EBM was a response to the belief that variation is caused by professional uncertainty. Likewise, quality of care benchmarks, and physician “report cards” have been advocated in an attempt to standardize practice under the belief that variation results in poor outcomes for some patients. The issue of variation has been an important one in the literature for at least four decades. The goal is implied: variation should be eliminated (Tanenbaum, 2013).

However, the methods used to identify variation and determine its sources, and the conclusions that have been based on these methods have gone relatively unchallenged. Throughout the course of the previous chapters I have identified a
number of issues with these methods. The first is the assumption that the residual variation in these studies is undesirable. An examination of the literature demonstrated that the field is inconsistent as to what constitutes unwarranted variation. Furthermore, those definitions that do exist are infrequently applied (explicitly). A clear and functional definition of unwarranted variation is both needed and is necessary to guide research so that only those variations that are considered undesirable can be teased out. The second issue is that variations studies have at best superficially accounted for systematic differences in the characteristics of populations, most notably related to the magnitude of healthcare needs. This can lead to erroneous conclusions regarding both the magnitude of (assumed unwarranted) variation and what factor(s) to which it is attributed.

Strategies to improve upon traditional methods, so that the results of variations studies may be more accurate in their conclusions and relevant to healthcare managers, were identified from the literature.

A common theme among variations studies is that physicians are an important cause of variation. Previous studies have not sufficiently isolated the effect of the physician as a source of variation from that due to patient and environmental factors. The presented study on variation between interventional cardiologists can be seen as an advancement on traditional methods in that the outcome of interest was the physicians’ recommendations rather than what the patient ultimately received in treatment. This is an important distinction. In contrast to the physician’s recommendation, what the patient ultimately receives is
not under the direct control of the physician. Another important contribution is that the study was better able to isolate variation at the physician level compared to previous studies. This was achieved by focusing on variation between physicians within a single centre, in a publicly funded healthcare environment, and where patients are referred to the centre rather than the physician. Although this study may have confirmed that variation is in part due to physicians, the mechanism(s) by which this variation arises remains unclear. The presented investigation into emergency medicine physicians’ management decisions suggests that patient social context is a contributing factor in determining the course of care. Social context has been overlooked in previous studies of practice variation. Recognition that this may be an important factor, and integration of social context into study methods would certainly advance our understanding of practice variation. Furthermore, the study is unique in that the “social context” was not defined explicitly by “objective” methods, but was inferred by the clinician.

7.2 Accounting for Health Care Needs

There is a notion that where different communities appear similar in their composition, the presence of variation indicates that some communities are using resources beyond what is necessary to achieve the observed health gain, and/or that some communities do not have access to the appropriate level of resources to achieve the desired health gains (Cohen et al., 1992; Bojakowski, 2010; Kappelman et al., 2007; Mulley, 2009). Interpreting variation in either of these ways can impact
how resources may be allocated. For example, in the event that variation is interpreted as “overuse” in relatively higher rate of use communities, healthcare managers may seek to restrict access or implement corrective mechanisms to ensure appropriate use. Likewise, where variation is interpreted as “underuse” in relatively lower rate of use communities, resources might be directed towards improving access (and thus, away from other programs). The appropriateness of these decisions is based on the extent to which observed variations are both real and problematic for meeting healthcare policy goals. Chapters 5 and 6 presented a number of issues that must be considered before responding to observed variations when planning healthcare services and allocating resources.

It is difficult to determine the extent to which variation is indeed indicative of over/underuse of healthcare services. In chapter 6 we saw that traditional methods for identifying variation may not be sensitive to population differences in healthcare need. If managers desire to allocate resources according to a policy of equal access/use for equal need, they must be careful not to confuse greater/lesser needs with over/underuse. Furthermore, managers need to discriminate between situations where underuse is due to lack of resources, and those due to “low” uptake of specific treatments despite available resources, as each of these relies on different solutions to minimize variation. Certainly, it would not be desirable to provide additional resources where sufficient resources exist and are not used.

How to best incorporate consideration of patient/population needs into studies examining variation requires more research. One approach is to directly
measure the needs for individuals in each population under investigation, and then adjust for systematic differences between populations when determining the rate of service use. Assuming one has the resources to do so, how one determines healthcare needs may be complex – there are varying degrees of severity for many disease states, with numerous combinations of co-morbidity and/or multi-morbidity that must be incorporated into a single measure of “need” for each individual. Future research should determine how to improve our ability to measure healthcare needs, with considerations of reliability and validity, and the feasibility of data collection for purpose of examining medical practice variations.

An alternative approach is to use indirect measures of need. Again, what constitutes an appropriate indirect measure of population need requires more research. Regardless of whether direct or indirect measures of need are used, the appropriate means to incorporate this information into analyses when examining medical practice variations must still be determined. This will require some consideration of the purpose of the study (e.g. to determine unwarranted variation – in that case, one needs a definition of what is “unwarranted”). While some methods for determining/accounting for needs were presented in the previous chapters, the extent to which these are appropriate is still to be determined.

7.3 Accounting for Patient Context

It is reasonable that an individual’s social context or social circumstances, such as employment status, social support, ability to access care, etc., can influence
the type of care that individual will ultimately receive. For example, an individual may be reluctant to undergo surgery, despite a demonstrated mortality benefit, if this surgery will result in financial hardship due to missed employment opportunities during recovery. Social context seems to be an important factor in determining how physicians use resources, as was demonstrated in the experiment examining emergency medicine physicians’ management decisions. This is not a novel concept. Models of healthcare services use from over four decades ago recognize the potential influence of social context in determining patient’s use of such services (e.g. Andersen and Newmen, 1973; Ro, 1969), and yet, despite the intuitive sense of the importance of social context, studies demonstrating its influence (e.g. Greenlick, 1970; Suchman, 1964; Dunlop et al., 1953), and the fact that experienced physicians seem to respond to this when determining how care will be managed, studies examining medical practice variations do not explicitly account for social context.

Why such studies overlook social context is not clear. It may be that as social context is multifaceted, it is functionally difficult to define, and thus measure. It may be that data regarding social context is not readily available to those researchers interested in examining medical practice variations. It may be, as in the emergency medicine experiment, that social context is highly idiosyncratic and case specific, and not easily captured in population level variables. Additional research is required to determine how to define, measure, and incorporate social context into medical practice variation studies. However, first we must confirm that social
context indeed influences management decisions in both a larger and more diverse physician population than that presented here, and in medical domains outside of emergency medicine. Furthermore, the extent to which a consideration of social context by the physician when determining the course of care actually improves patient important outcomes requires further investigation. That is, we require a deeper understanding of when and how social context should be incorporated into management decisions before we can determine if variation due to differences in social context is reasonable.

7.4 The Physician as a Source of Variation

The idea that physicians are a significant contributor to observed variations has received attention in the published literature over the past eight decades. While it is a reasonable notion, conclusively demonstrating this has been difficult. One reason is the emphasis on outcome measures that describe how the patient was ultimately managed, rather than what the physician intended. Certainly, there are a number of factors that can affect how the patient is managed beyond what the physician would recommend. Such factors include patient preferences, resource availability, affordability, and (as introduced in a previous chapter) perhaps social context. The presented study of interventional cardiologists minimized the influence of such factors, focusing on decisions rather than outcomes in an environment common to each observed cardiologist. Thus, it would seem that physicians are in some part a source of variation.
What is not clear is why physicians might differ in how they manage similar patients. Some suggested explanations include differing levels of knowledge, reactions to professional uncertainty, and income preferences. While these factors cannot be ruled out (and are certainly, to some extent, reasonable/likely explanations), the available data is by no means conclusive on the absolute or relative influence of each. The study of emergency medicine physicians introduced the possibility of patient social context as a source of variation. Social context was not measured in previous studies demonstrating physicians as a source of variation, including the above-mentioned study of interventional cardiologists. Thus, one cannot rule out the possibility of systematic differences in the social context of patients between observed physicians. However, it is unlikely that patients in the interventional cardiology study systematically differed among observed cardiologists to the extent that there is such a magnitude of difference in recommended treatment, especially considering the pseudo-random nature of patient assignment and the extent of adjustment for important clinical and demographic characteristics. This might hint at the possibility that physicians differently react to or solicit information regarding social context – a subject for future study.

One aspect of the management decision that has been overlooked in practice variations studies implicating the physician as an important contributor is perception. Advances in cognitive science demonstrate that even when individuals are presented with the same information, features of perception can bias how that
information will be interpreted, and thus, the decisions those individuals make. For example, cases/events that are more easily recalled or retrieved from memory (perhaps because they frequently occur, or were dramatic in nature) are often given too much weight by the individual (i.e. “availability bias”) (Tversky and Kahneman, 1974). Likewise, we know from previous research that a decision maker’s assessment of risk and benefit for a given choice option can be affected by positive or negative feelings toward some feature of the option (i.e. “affect bias”) (Finucane, Alhakami, Slovic, & Johnson, 2000; Slovic, Finucane, Peters, & MacGregor, 2004). Thus, it may be that physicians are similar in how they approach the clinical problem but differ in their execution (of the management decision) because they differ in how they perceive clinical information, perhaps due to experience, contextual features, or recent/memorable cases (Brooks, LeBlanc, & Norman, 2000). Variations studies might benefit from incorporating methods from cognitive psychology and/or decision-making science to reveal mechanisms by which individual physicians might differ in their management decisions. This may help to improve strategies to minimize unwarranted variation.

7.5 Overall Picture

In the event that one is interested in describing variation that is a direct result of a particular factor, then it is important to tease out all variation that is caused by other factors. This requires complete knowledge of which factors cause variation. Previous studies have overlooked many important factors when
determining the existence and sources of variation. Thus, a significant portion of the variation remained unexplained. This is problematic because in most cases the unexplained variation was assigned to one unmeasured factor or another, from which corrective mechanisms to eliminate variation were advocated. For example, when demographics or available resources did not explain variation, physicians’ “practice style” was deemed the culprit (Wennberg, 1984). This was assumed to be a function of their level of knowledge, and thus, clinical studies and knowledge translation strategies were advocated. In the event that variation existed for good reason, or was due to a factor that was not considered, such strategies to eliminate it would be unsuccessful, or might even result in suboptimal patient outcomes.

Why some factors may have been overlooked may be due to a lack of access to necessary data. However, it may also be due to poor knowledge of researchers regarding which factors are important. An examination of the existing literature (Chapters 2 and 6) suggests that differences in patient health status and needs are causes of variation, as are differences in available resources and access to them. The presented study of interventional cardiologists (Chapter 3) demonstrates that a significant portion of the variation remains unexplained, even when patient health status is considered, and factors related to accessibility and environment are controlled, indicating that additional sources of variation should be considered. The study of emergency department physicians (Chapter 4) suggests that patient social context may play a role in determining how a patient’s care is ultimately be managed. As described above, factors operating at the level of a physician’s decision
making process (e.g. perception, heuristics, and biases) may be potential sources of variation. Currently, there are no theories or frameworks for studying practice variation that recognize each of these factors. Such a theory is important for a number of reasons. Most important among these reasons is that theory will help researchers develop their studies to include all the relevant factors, such that a significant portion of the variation can be explained.

While additional work is needed to develop a theory, such a theory should include three general factors: 1) the patient, 2) the physician, and 3) the environment. Although these general factors have been considered in the published literature, what the studies presented in previous chapters reveal is that each of these factors should be broken down further. For example, when considering the patient, one should bear in mind health status, ability to access care, affordability, preferences, and social context. One should consider not only characteristics of the physician’s training and preferences, but also how he or she approaches the management problem. Consideration of the environment should not be restricted to available resources at the institution where care is delivered (e.g. available beds), but also to relevant resources in the community.

From the perspective of improving both quality of care and resource allocation, it is the unwarranted variations that are of interest. A theory may include consideration of unwarranted variation. However, the argument presented in Chapter 5 indicates that what constitutes unwarranted variation requires more examination. Ultimately, what one considers as unwarranted variation may be
value-laden. Application of a comprehensive theory of practice variation when performing a study should yield a high amount of explained variance attributable to defined factors. If the relative influence of each factor is known, and the reader has defined which factors are undesirable, one can determine the proportion of the variation that is unwarranted. Only then can appropriate strategies be developed to ensure that only that variation that does not exist for good reason is eliminated.

**7.6 Conclusions**

The value of medical practice variations studies is contingent on what we can learn from the data they present. Thus far, the data has been incomplete. What is most distressing is the lack of theory or a comprehensive framework to guide these studies. As a result, variations studies have focused on identifying variation and then determining the cause post hoc. Many of the issues presented could have been mitigated if researchers had a clear picture of what to look for in variation prior to seeking out its existence. Consideration of the issues presented above and in previous chapters should lead to improved methods for examining medical practice variation. Most important, the previous chapters speak to the need for theory driven research in medical practice variations. Only in this way can we devise studies to truly understand if and why similar patients differ in the care they receive.
References


