INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®
PHARMACEUTICAL GUIDELINE COMPLIANCE AND ITS IMPACT ON COSTS AND EFFECTIVENESS:
Case Studies of orders based on Vancomycin Use and Intravenous to Oral Switch
Antimicrobial Guidelines at The Ohio State University Medical Center

DISTRIBUTION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate School
of The Ohio State University

By:

Dennis W. Grauer, M.S.

*****

The Ohio State University

2000

Dissertation Committee:
Dev S. Pathak, D.B.A., Adviser
Mitchell Medow, M.D.
Craig A. Pedersen, Ph.D.

Approved by

Dev S. Pathak, D.B.A., Adviser
College of Pharmacy
ABSTRACT

The goals of practice guidelines are to improve quality and reduce cost. However, evaluation of guidelines relative to goals has been limited. A paucity of studies on locally developed guidelines is found in the literature. The purpose of this study was to examine the effect of guidelines on outcomes. The guidelines studied were Converting From Intravenous to Oral Antimicrobial Therapy and Criteria for Use of Vancomycin. The guidelines were examined independently and the study design was a two group, observational, cohort design. Groups were defined based on treatment compliant with guideline recommendations versus not compliant. The target populations were patients with drug orders for medications covered by the guidelines. Data were collected via written baseline and follow-up survey, chart review and hospital database. The first three objectives were to evaluate the effect of compliance on economic, clinical and health status outcomes. The final objective was to determine if order compliance was cost effective relative to non-compliance. For the vancomycin guideline, the study determined compliance did not have statistically significant effect on economic or health status outcomes. However, the results were significant since order compliance produced clinical outcomes that were better than non-compliant with the guideline. Specifically, mortality was lower and antimicrobial success rates were
higher for the compliant cohort. Despite improved clinical effect, the study did not find
evidence that following guideline recommendations for the vancomycin guideline was
more cost effective. Similarly, for the iv to po guideline, statistically significant effects
of order compliance was not found for economic, clinical, health status outcomes or on
cost effectiveness analysis. However, the non-compliant cohort had higher total
hospital costs and lower lengths of stay compared to the compliant cohort. This result
indicates that physicians were more likely to minimize iv to po switch due to potential
drug failure error (Type I error) and instead keep patients on iv drug therapy longer
even though the switch to po may have been indicated by the guideline (Type II error).
Evaluation protocols for guidelines should be examined and specified based on the
goals of the guidelines prior to implementation.
Dedicated to my wife, Sharlyn

And

Megan and Tyler, my children
ACKNOWLEDGMENTS

I would like to express sincere gratitude to the following individuals whose contributions to the completion of this work were instrumental.

Dev Pathak for his counsel and support on this and other projects, for his guidance throughout the entirety of my educational endeavors and for his persistence and aid in my finishing this portion of my education.

Craig Pedersen for serving as counsel for the project and for being a significant source of testing my ideas associated with this project.

Mitchell Medow for aiding clinical interpretation of this project and for being the primary clinical resource for the project.

John Sauer for helping with the collection and gathering of much of the data for this study.

Karen Hale for assistance in getting the project ready for and submitted for Investigational Review.

Jerry Seigel, Lou Mavko, administrative staff and all residents in the Department of Pharmacy who provided significant aid in identifying the patients for this study and who had a large impact on my past educational and professional endeavors.

My fellow graduate students whom allowed me to vent and share stories with and who provided constant feedback and aid whenever asked.

Hagop Mekhjian, Gail Marsh and other hospital employees for their support and education with respect to the topic of this project and many other health care issues.

The Ohio State University Medical Center for allowing the study of the two guidelines which are copyrighted by the Ohio State University and for permission to publish the guidelines in this report. All Rights Reserved.

The Medical Outcomes Trust for granting permission to use and reproduce the SF-12™ Health Survey, Copyright © 1994 The Health Institute; New England Medical Center. All Rights Reserved.
Sue Keyser and Kathy Sloan for administrative assistance.

Kathy Brooks for helping one of her problem graduate students and making life for graduate students easier in the College of Pharmacy.

Ruth and Norman Grauer and Dianne and Erwin Fink, my parents and parents-in-law, for providing significant aid and support of my family while this project was finished.

Dave Grauer, my brother, for consistent career advice and for significant aid and support throughout my stay.

Phyllis, Allison, and Kevin Grauer, my brother's family, for sharing time with me when I needed rest, food and an outlet for entertainment.

Sharlyn Grauer, for happily supporting me throughout my time spent at Ohio State, for her significant role in getting me to this point and for her unending aid in all that I do.

Megan Grauer and Tyler Grauer, my daughter and son, for providing constant joy and motivation.

Pharmacia & Upjohn, Inc. whose financial support was greatly appreciated and resulted in the completion of the project.
VITA

October 16, 1964  
Born:  
Marysville, Kansas

May 1989  
B.S. Pharmacy:  
The University of Kansas

Pharmacy Practice Resident:  
The Ohio State University Medical Center

June 1991  
M.S. Hospital Pharmacy:  
The Ohio State University

Graduate Teaching Associate:  
The Ohio State University

July 1995 - August 1996  
Syntex Post Graduate Pharmacoeconomic Fellow:  
Division of Transplantation, The Ohio State University Medical Center

August 1996 – October 1998  
Director, Outcomes Measurement & Research:  
The Ohio State University Medical Center

PUBLICATIONS


FIELDS OF STUDY

Major Field: Pharmacy

Secondary Field: Marketing
# TABLE OF CONTENTS

Abstract................................................................................................................ ii

Dedication............................................................................................................ iv

Acknowledgments............................................................................................... v

Vita....................................................................................................................... vii

List of Tables...................................................................................................... xv

List of Figures...................................................................................................... xvii

Chapters:

1. Introduction.................................................................................................... 1

   Background.................................................................................................... 2
   Interest In Clinical Practice Guidelines...................................................... 3
   Definition of Clinical Practice Guideline.................................................... 4
   Characteristics of ‘Good’ Practice Guidelines............................................ 5
   Use of Clinical Practice Guidelines............................................................ 8

Context of the Study...................................................................................... 10
   Guideline Development at OSUMC.............................................................. 11
   Locally Developed Practice Guidelines..................................................... 12
   OSUMC Practice Guidelines....................................................................... 13

Impact of Practice Guidelines on Outcomes.................................................. 14

Research Questions......................................................................................... 15

   Impact of Demographic Variables........................................................... 16
   Impact of Clinical History.......................................................................... 17
   Guideline Compliance (Process Variables)................................................ 18

Economic Outcome......................................................................................... 20

   Research Question One............................................................................ 22

Clinical Outcome............................................................................................ 23

   Research Question Two............................................................................. 24
2. Literature Review .................................................................................. 41

Practice Guideline Process........................................................................ 42
Development................................................................................................ 42
   Topic Selection (Step 1)......................................................................... 44
   Evidence Gathering and Assimilation (Steps 2 and 3).......................... 45
   Population Definition (Step 4).............................................................. 49
   External Review (Step 5)....................................................................... 49
   Adoption of the Practice Guideline (Steps 6 and 7)............................ 50
   Review and Update of the Practice Guideline (Step 8)....................... 50
   National Versus Local Development.................................................... 50
   Barriers to Development....................................................................... 52
   Practice Guideline Development Quality............................................. 53
Implementation (Education)......................................................................... 54
   Methods of Implementation.................................................................... 55
   Medical Information Diffusion Models................................................. 56
   Studies of Implementation (Compliance).............................................. 58
   Barriers to Implementation (Compliance)............................................... 64
Evaluation.................................................................................................. 67
   Methodology of Evaluation.................................................................... 71
   Attitudes................................................................................................. 71
   Knowledge............................................................................................. 73
   Behavior............................................................................................... 74
   Behavior – Observational Studies......................................................... 75
   Behavior – Pre and Post Implementation Studies.................................. 76
   Behavior – Interventional Studies.......................................................... 80
   Evaluation (Outcomes)......................................................................... 81
   Clinical Outcome................................................................................... 81
   Economic Outcome............................................................................... 84
   Health Status and Cost Effectiveness.................................................... 86
   Evaluation Barriers................................................................................ 87
   Evaluation Summary............................................................................. 88
   Practice Guideline Literature Summary.............................................. 88
   General Concerns with Practice Guidelines......................................... 90
3. Methods....................................................................................................... 129

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Design</td>
<td>130</td>
</tr>
<tr>
<td>Study Treatment</td>
<td>132</td>
</tr>
<tr>
<td>Data Collection</td>
<td>132</td>
</tr>
<tr>
<td>Patient Confidentiality</td>
<td>135</td>
</tr>
<tr>
<td>Guideline Evaluation</td>
<td>136</td>
</tr>
<tr>
<td>Nonresponse and Response Rate</td>
<td>136</td>
</tr>
<tr>
<td>Eligible versus Not Eligible</td>
<td>137</td>
</tr>
<tr>
<td>Enrolled Patients versus Not Enrolled Patients</td>
<td>138</td>
</tr>
<tr>
<td>Non Response Analysis</td>
<td>138</td>
</tr>
<tr>
<td>Lost To Follow-Up</td>
<td>139</td>
</tr>
<tr>
<td>Item Non Response</td>
<td>139</td>
</tr>
<tr>
<td>Response Rate</td>
<td>140</td>
</tr>
<tr>
<td>Instrument Development</td>
<td>140</td>
</tr>
<tr>
<td>Demographics</td>
<td>140</td>
</tr>
<tr>
<td>Clinical History</td>
<td>141</td>
</tr>
<tr>
<td>Guideline Compliance or Process Variables</td>
<td>142</td>
</tr>
<tr>
<td>Economic Outcomes</td>
<td>142</td>
</tr>
<tr>
<td>Clinical Outcomes</td>
<td>142</td>
</tr>
<tr>
<td>Health Status</td>
<td>143</td>
</tr>
<tr>
<td>Cost Effectiveness Analysis (CEA)</td>
<td>143</td>
</tr>
<tr>
<td>Instrument and Tool Evaluation Procedures</td>
<td>143</td>
</tr>
<tr>
<td>Panel of Experts</td>
<td>144</td>
</tr>
<tr>
<td>Field Test</td>
<td>144</td>
</tr>
<tr>
<td>Pilot Study</td>
<td>144</td>
</tr>
<tr>
<td>Instrument Validity</td>
<td>145</td>
</tr>
<tr>
<td>Criterion and Criterion-Related Validity</td>
<td>145</td>
</tr>
<tr>
<td>Content Validity</td>
<td>146</td>
</tr>
<tr>
<td>Construct Validity</td>
<td>146</td>
</tr>
<tr>
<td>Reliability/Health Status Instrument</td>
<td>147</td>
</tr>
<tr>
<td>Reliability/Chart Review Form</td>
<td>149</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>150</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>Target Population and Sampling</td>
<td>151</td>
</tr>
<tr>
<td>Target Population</td>
<td>151</td>
</tr>
<tr>
<td>Sampling Procedures/Data Collection</td>
<td>152</td>
</tr>
<tr>
<td>Sample Size</td>
<td>154</td>
</tr>
<tr>
<td>Statistical Analysis</td>
<td>157</td>
</tr>
<tr>
<td>Independent Samples t-test</td>
<td>158</td>
</tr>
<tr>
<td>Chi-Square Test of Independence</td>
<td>159</td>
</tr>
<tr>
<td>Research Question One</td>
<td>162</td>
</tr>
<tr>
<td>Analysis of Covariance</td>
<td>163</td>
</tr>
<tr>
<td>Research Question Two</td>
<td>164</td>
</tr>
<tr>
<td>Research Question Three</td>
<td>165</td>
</tr>
<tr>
<td>Research Question Four</td>
<td>167</td>
</tr>
<tr>
<td>Design Limitations</td>
<td>168</td>
</tr>
<tr>
<td>History</td>
<td>168</td>
</tr>
<tr>
<td>Testing</td>
<td>168</td>
</tr>
<tr>
<td>Statistical Regression</td>
<td>169</td>
</tr>
<tr>
<td>Selection</td>
<td>169</td>
</tr>
<tr>
<td>Selection-Maturation Interaction</td>
<td>169</td>
</tr>
<tr>
<td>External Validity</td>
<td>169</td>
</tr>
<tr>
<td>The Reactive or Interaction Effect of Testing</td>
<td>170</td>
</tr>
<tr>
<td>Interaction of Selection Bias and the Experimental Variable</td>
<td>170</td>
</tr>
<tr>
<td>Reactive Effects of the Experimental Arrangements</td>
<td>170</td>
</tr>
<tr>
<td>References for Chapter 3</td>
<td>171</td>
</tr>
<tr>
<td>4. Results</td>
<td>173</td>
</tr>
<tr>
<td>Guideline Evaluation</td>
<td>174</td>
</tr>
<tr>
<td>Psychometric Properties of Instruments</td>
<td>176</td>
</tr>
<tr>
<td>Validity</td>
<td>176</td>
</tr>
<tr>
<td>Reliability – Health Status</td>
<td>178</td>
</tr>
<tr>
<td>Reliability – Chart Review</td>
<td>179</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>180</td>
</tr>
<tr>
<td>Response Rate</td>
<td>182</td>
</tr>
<tr>
<td>Organization of the Remainder of Results</td>
<td>187</td>
</tr>
<tr>
<td>Vancomycin Guideline</td>
<td>187</td>
</tr>
<tr>
<td>Demographic Characteristics</td>
<td>187</td>
</tr>
<tr>
<td>Clinical History Characteristics</td>
<td>190</td>
</tr>
<tr>
<td>Economic Outcome (Research Question One)</td>
<td>191</td>
</tr>
<tr>
<td>Clinical Outcome (Research Question Two)</td>
<td>193</td>
</tr>
<tr>
<td>Health Status Outcome (Research Question Three)</td>
<td>193</td>
</tr>
<tr>
<td>Cost Effectiveness (Research Question Four)</td>
<td>195</td>
</tr>
<tr>
<td>Vancomycin Guideline Conclusion</td>
<td>196</td>
</tr>
<tr>
<td>IV to PO Guideline</td>
<td>196</td>
</tr>
<tr>
<td>Demographic Characteristics</td>
<td>196</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Clinical Practice Guideline Standards</td>
<td>7</td>
</tr>
<tr>
<td>2.1</td>
<td>Steps to Develop Practice Guidelines</td>
<td>44</td>
</tr>
<tr>
<td>2.2</td>
<td>Full Economic Evaluation Techniques</td>
<td>107</td>
</tr>
<tr>
<td>3.1</td>
<td>Medication Therapies Covered By Guidelines</td>
<td>132</td>
</tr>
<tr>
<td>4.1</td>
<td>OSUMC Guideline Development Standards</td>
<td>175</td>
</tr>
<tr>
<td>4.2</td>
<td>Validity Correlations - PCS-12 By Number of Comorbid Conditions</td>
<td>177</td>
</tr>
<tr>
<td>4.3</td>
<td>Validity Correlations - MCS-12 By Number of Comorbid Conditions</td>
<td>177</td>
</tr>
<tr>
<td>4.4</td>
<td>Comparison of Mean MCS-12 Scores By Mental Illness ICD-9 Code</td>
<td>177</td>
</tr>
<tr>
<td>4.5</td>
<td>‘Ceiling’ and ‘Floor Effects of the SF-12 Subscales</td>
<td>179</td>
</tr>
<tr>
<td>4.6</td>
<td>Chart Review Agreement (Reliability)</td>
<td>180</td>
</tr>
<tr>
<td>4.7</td>
<td>PCS-12 Difference Score As Assessed by Global Ratings</td>
<td>181</td>
</tr>
<tr>
<td>4.8</td>
<td>MCS-12 Difference Score As Assessed by Global Ratings</td>
<td>181</td>
</tr>
<tr>
<td>4.9</td>
<td>PCS-12 Difference Score Tests of Between Subjects Effects</td>
<td>182</td>
</tr>
<tr>
<td>4.10</td>
<td>MCS-12 Difference Score Tests of Between Subjects Effects</td>
<td>182</td>
</tr>
<tr>
<td>4.11</td>
<td>Response Analysis - Vancomycin</td>
<td>185</td>
</tr>
<tr>
<td>4.12</td>
<td>Response Analysis - IV to PO</td>
<td>186</td>
</tr>
<tr>
<td>4.13</td>
<td>Demographic Comparison of CGP and NCGP - Vancomycin</td>
<td>189</td>
</tr>
<tr>
<td>4.14</td>
<td>Clinical History Characteristics - Vancomycin</td>
<td>190</td>
</tr>
<tr>
<td>4.15</td>
<td>Economic Outcome - Vancomycin</td>
<td>192</td>
</tr>
<tr>
<td>4.16</td>
<td>Clinical Outcome - Vancomycin</td>
<td>193</td>
</tr>
<tr>
<td>4.17</td>
<td>Follow-Up Health Status Outcome - Vancomycin</td>
<td>194</td>
</tr>
<tr>
<td>4.18</td>
<td>Vancomycin PCS-12 Difference Score Descriptives</td>
<td>194</td>
</tr>
<tr>
<td>4.19</td>
<td>Vancomycin MCS-12 Difference Score Descriptive</td>
<td>194</td>
</tr>
<tr>
<td>4.20</td>
<td>PCS-12 Difference Score Tests of Between Subjects Effects (ANCOVA) - Vancomycin</td>
<td>194</td>
</tr>
<tr>
<td>4.21</td>
<td>MCS-12 Difference Score Tests of Between Subjects Effects (ANCOVA) - Vancomycin</td>
<td>195</td>
</tr>
<tr>
<td>4.22</td>
<td>Cost Effectiveness - Vancomycin Guideline</td>
<td>196</td>
</tr>
<tr>
<td>4.23</td>
<td>Demographic Comparison of CGP and NCGP - IV to PO</td>
<td>197</td>
</tr>
<tr>
<td>4.24</td>
<td>Clinical History Characteristics - IV to PO</td>
<td>198</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.25</td>
<td>Economic Outcome – IV to PO</td>
<td>200</td>
</tr>
<tr>
<td>4.26</td>
<td>Clinical Outcome – IV to PO</td>
<td>201</td>
</tr>
<tr>
<td>4.27</td>
<td>Follow-Up Health Status Outcome – IV to PO</td>
<td>202</td>
</tr>
<tr>
<td>4.28</td>
<td>IV to PO PCS-12 Difference Score Descriptives</td>
<td>202</td>
</tr>
<tr>
<td>4.29</td>
<td>IV to PO MCS-12 Difference Score Descriptive</td>
<td>202</td>
</tr>
<tr>
<td>4.30</td>
<td>PCS-12 Difference Score Tests of Between Subjects Effects (ANCOVA) – IV to PO</td>
<td>202</td>
</tr>
<tr>
<td>4.31</td>
<td>MCS-12 Difference Score Tests of Between Subjects Effects (ANCOVA) – IV to PO</td>
<td>203</td>
</tr>
<tr>
<td>4.32</td>
<td>Cost Effectiveness – IV to PO Guideline</td>
<td>203</td>
</tr>
<tr>
<td>5.1</td>
<td>Development Standards Met – OSUMC versus Published Results</td>
<td>208</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure

1.1 The Clinical Practice Guideline Process ................................................. 10
2.1 Attitude-Knowledge-Behavior Paradigm ................................................ 70
3.1 Data Collection Flow Diagram ................................................................. 135
5.1 Guideline Compliance Decision Making ................................................ 220
CHAPTER 1
INTRODUCTION

The purpose of this dissertation is to examine the cost and effectiveness of drug therapy order compliance with the recommendations made by two locally developed and implemented drug therapy related clinical practice guidelines. Specifically, the study will answer the following question, “What is the cost and effectiveness of compliance with locally developed drug-therapy practice guidelines”? Current growth in the development of clinical practice guidelines has not been associated with a corresponding growth in evaluation of the effect of the guideline on the outcomes targeted by the respective guideline. This chapter first provides background information on clinical practice guidelines. Then the context of the study is reviewed followed by a discussion on the potential effect of drug therapy clinical practice guidelines. The chapter continues with an explanation of the variables used and the research questions posed by the study. Then, assumptions and limitations of the study are noted, and definitions of key terms are provided. The chapter ends with a brief review of the organization of the rest of the dissertation.
BACKGROUND

In the past sixty years, medical advancement has been largely due to increased understanding of disease processes, improved diagnostic ability and improved treatment options. It is no surprise that the majority of research in health care during this period focused on new medical discovery.

Economic pressure felt by today's health care system escalated in the late 1980's. Health care continues to claim an increasing share of society's human, financial and physical resources. Five major factors responsible for increased health care spending include: increasing population size, increasing elderly population, rising general inflation, growing medical care price inflation and increasing volume and intensity of services. Increased resource utilization has erected a significant barrier to improving the access and quality of health care for the population as a whole. Current economic pressure demands the health care delivery system be fiscally accountable. Yet, not only should the health care delivery system be economically sound but effectiveness must at least be maintained or even improved. Attempts to control health care expenditure risk curtailing necessary as well as unnecessary health services. It has been shown that patients cannot discriminate between appropriate versus inappropriate care and tend to reduce both when economic incentives are applied. The challenge is to reduce health services that are of little or no benefit to patients, while permitting and encouraging the delivery of necessary
care. The accountability issue increased need and interest on research of how the process of health care delivery impacts patient care. Demand for accountability has aided the proliferation of clinical practice guidelines.

**Interest In Clinical Practice Guidelines**

Gottlieb et al. stated the reasons for heightened interest in practice guidelines include creating an economy of care, reducing variation in practice, maintaining basic levels of care, and improving the quality of care. Berg et al. states that interest in practice guidelines is a result of physicians inability to keep up with current literature, the need to eliminate implicit in-correct guidelines, and to increase patient participation in decision processes. These are all valid explanations of the interest in practice guideline development. Growth in development continues despite the relative paucity of literature evaluating the effect of guidelines. Research clearly indicates variability occurs in the health care delivery process and creates uncertain patient outcomes.

Studies demonstrate that there are significant differences in the geographic utilization rates of common and expensive surgeries such as hysterectomies, tonsillectomy and prostatectomy. Thus, patients with similar conditions may be treated differently or have different outcomes depending on geographic location, type of provider, local providers and ethnic groups. One conclusion from these studies is that there is mis-utilization (i.e., either over-utilization or under-utilization) of health
based on factors external to the individual patient. Often it is assumed the outcome and/or process is different based on these factors. An extension of this rationale holds that variation in processes of health care delivery does not lead to optimal final patient outcomes. In fact, variation in practice has been observed to be the greatest when providers are uncertain about the most appropriate approaches to care. The logic further assumes that patients with the same condition have no other differences which may influence patient care decisions or outcomes. Also underlying the assumption is that alternative therapies are available which produce more positive outcomes. The result is that variation has been used as a crude marker for 'negative' quality in the health care delivery process. However, this variation does not necessarily mean a variation in quality. In fact, the absence of strong literature evidence may indicate that some variation in practice is appropriate.

**Definition of Clinical Practice Guidelines**

The development of clinical practice guidelines is one attempt to standardize and reduce variation in the health care delivery process. Practice guidelines are an attempt to discourage ineffective medical practice, encourage effective practice and improve health outcomes. At the national level, The Agency for Health Care Policy and Research (AHCPR), formed in 1989 and was given the charge by The United States Congress to facilitate, disseminate and evaluate clinical practice guidelines. The assumption again is that reducing variation will lead to an increase in the quality of care.
delivered to patients, a reduction in the cost of health care delivered or perhaps both. Other uses of practice guidelines will be discussed later in this paper.

In general, practice guidelines are tools designed to help practitioners, payers and patients reduce the rate of rising costs while not sacrificing quality.\(^1\) The Institute of Medicine (IOM) defines practice guidelines as "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances".\(^13\) Eddy et al. defined practice guidelines as "a preformed recommendation made for the purpose of influencing a decision about a health intervention".\(^14\) The key term used in each definition and the focus of this study is the word decision. Practice guidelines are to be used in some way by clinicians and patients when making decisions. Clinical practice guidelines are referred to in the literature by various names including clinical pathways, practice parameters, patient care paths, practice policies, clinical pathways, standards and algorithms.\(^4\)\(^15\) For the purpose of this study, the only term used will be practice guideline or clinical practice guideline or guideline.

**Characteristics of ‘Good’ Clinical Practice Guidelines**

Lohr (1994) provided general criteria for ‘good’ practice guidelines.\(^16\) The assumption is that developing guidelines within this list of attributes will produce practical, explicit, working documents, not just healthy compilations of the literature.
The criteria for good guidelines state that they should be:

1. valid
2. reliable and reproducible
3. clinically applicable
4. clinically flexible
5. clear
6. produced with multidisciplinary input
7. reviewed periodically
8. document able

The American Medical Association (AMA), the Institute of Medicine (IOM) and the Canadian Medical Association (CMA) have each formulated independent methodology for developing scientifically rigorous guidelines.\textsuperscript{13,14,18-20} Based on these published development methodology criteria, Shaneyfelt et al. developed and tested an observational instrument that measures adherence to the suggested methodologies when developing a practice guideline\textsuperscript{20}. The list of attributes appears in Table 1.1 and will be used later in this study to judge the relative merits of the guidelines in this study. Thus, the merits of practice guidelines may be judged relative to the standards established by leading professional associations and the literature. A clinical practice guideline may be developed to meet these published criteria. In fact, following the criteria may improve the face and content validity of the clinical practice guideline document. Yet, meeting all of the criteria does not provide any information with respect to the effect of the practice guideline after the guideline has been implemented.
**Standards of Guideline Development**

1. Purpose of the guideline is specified
2. Rationale and importance of the guideline are explained
3. The participants in the guideline development process and their areas of expertise are specified
4. Targeted health problem or technology is clearly defined
5. Targeted patient population is specified
6. Intended audience or users of the guideline are specified
7. The principal preventive, diagnostic, or therapeutic options available to clinicians and patients are specified
8. The health outcomes are specified
9. The method by which the guideline underwent external review is specified
10. An expiration date or date of scheduled review is specified
11. Method of identifying scientific evidence is specified
12. Time period from which evidence is reviewed is specified
13. The evidence used is identified by citation and reference
14. Method of data extraction is specified
15. Method for grading or classifying the scientific evidence is specified
16. Formal methods of combining evidence or expert opinion are used and described
17. Benefits and harms of specific health practices are specified
18. Benefits and harms are quantified
19. The effect on health care costs from specific health practices is specified
20. Costs are quantified
21. The role of value judgments used by the guideline developers in making recommendations is discussed
22. The role of patient preferences is discussed
23. Recommendations are specific and apply to the stated goals of the guideline
24. Recommendations are graded according to the strength of the evidence
25. Flexibility in the recommendations is specified

**TABLE 1.1: Clinical Practice Guideline Standards**
Uses of Clinical Practice Guidelines

Traditional quality measurement in health care relied on the identification of single instances of poor quality through case-by-case review. The process of case review itself relied on a single physician arbitrator using implicit criteria to classify a case as acceptable or unacceptable. Institutional, payer and judicial systems supported and continue to support the use of quality measurement as judged by a single case reviewer. The obvious problem with this approach is random error or variation. The same case may be evaluated and classified differently by two independent reviewers. This form of quality measurement relies solely on the expert opinion of a single individual. Case review has been described in the literature as ineffective, inefficient and antagonistic.

To improve the objectivity of quality measurement, health care professionals were asked to document explicit review criteria and cases were reviewed when it was judged to fail written criteria. Quality measurement needed structure to apply explicit criteria a-priori to measure quality. One method used to make criteria explicit was the development of clinical practice guidelines. Clinical practice guidelines were published and provided instant, explicit criteria for quality evaluation. Practice guidelines became standards of health care delivery upon which to judge practitioner behavior.

Health care standards span a spectrum from statutes and regulations to practice guidelines. Currently, health care standards address all aspects of health care including the availability, safety, efficacy, quality and cost of products and services. A clear trend
has been to move away from the more compulsory standard, such as a regulation, to a more discretionary standard such as a practice guideline.24 The former are set by official public agencies while the latter are standards set by quasi-official and voluntary bodies. Thus provider associations, payers, and provider organizations are using practice guidelines to set standards of care in an explicit manner. However, it must be kept in mind that practice guidelines codify standards of care but do not establish that standard. Actual practice dictates the standard.

Each practice guideline has a central purpose(s). That purpose may include the reasons for heightened interest listed above. Furthermore, the purpose should be spelled out in the guideline as the standards presented in Table 1.1 suggest. The following list has been suggested as practical uses of practice guidelines in addition to helping physicians make good decisions:16

1. pre-certification screening,
2. retrospective (utilization) review,
3. determination of justifiable payment, and
4. credentialing of physicians for permission to perform certain procedures.

As suggested by this list, current uses of practice guidelines are highly controversial. Uses of practice guidelines may be classified from the clinician point of view as positive or negative.4 Positive uses include making better clinical decisions, improved education, better quality improvement and more cost accountability. Negative uses include increased malpractice lawsuits and protecting clinical turf. Thus, depending on perspective, many uses of guidelines may appear to be either a threat or an opportunity.
The primary goal is for each practice guideline to have a central purpose which is specified and attainable. The group or individuals developing a guideline do so under the auspices of reaching a particular goal. This fact raises concern regarding biases introduced into the guideline by the developers. One result may be the decision not to comply with a guideline or part of a recommendation.

CONTEXT OF THE STUDY

The guideline process as defined in this study is presented in Figure 1.1. The process includes development, implementation and evaluation of the practice guideline. This dissertation focuses on practice guideline evaluation, but should not be considered completely independent from guideline development and implementation. The development and implementation processes have been shown to impact the actual use of a guideline in practice.\[4,25-28\]

![Figure 1.1: The Clinical Practice Guideline Process](image)
Guideline Development at The Ohio State University (OSUMC)

At OSUMC, a clinical governance committee has the charge to develop and approve practice guidelines. Topics are discussed and selected by the Committee or individual practitioners based on literature and clinical areas where improvement potential is felt to exist. Explicit criteria are used to determine the topic of a guideline. The Committee meets to discuss data and evidence prior to selecting a topic for which a guideline might be useful for development. However, the result is a large quantity and variety of guidelines with an equal variety of goals. In the next phase of the process, the topics are assigned to in-house clinical experts. The experts are responsible for conducting a literature review and utilizing expert opinion to develop a draft practice guideline. The draft practice guideline is submitted to internal and external, independent reviewers. Reviewer comments and corrections are made and the revised version of the guideline is again presented to the governance committee for final approval. If the guideline is approved, hard copies are distributed to all physicians and medical units and the guideline is placed on an internal computer network. If the guideline is not approved, it is returned to the expert for corrective action. Thirty-six clinical practice guidelines have been approved and sixty-four are in various stages of development. The main characteristics of the process used at OSUMC are that the guidelines are developed locally by an informal consensus approach and implemented with printed materials. These two characteristics are covered in detail in the Literature Review section of this dissertation.
Locally Developed Practice Guidelines

A national guideline developed by Middleton, et al. provides a valuable perspective on developing a national practice guideline. The guideline developed was titled, 'National Cancer Guideline for Prostate Cancer'. A major limitation to a national process is the need to find consensus. It has already been established in this paper that medical care can vary by geographic location. Thus, it should not be expected to get a consensus recommendation at a national level. The 'National Cancer Guideline for Prostate Cancer' made few recommendations. The guideline produced a summary of the current literature. Even the authors conclude that "practitioners should be aware of all (treatment) options". Many national guidelines are vague in both language and recommendations. Shapiro et al. concluded that to have effective practice guidelines, it is important that current practice and biases are not simply codified.1

The response of many local organizations is to modify national guidelines or develop local guidelines to make them more accommodating to local clinical practice. Local development also allows the local organization to provide more specific clinical recommendations. Brown et al. locally modified and adopted the national AHCPR Depression guideline. The result was a document which was smaller and more user-friendly. However, the effectiveness of the modified guideline was not reported. Grilli et al. cautioned on the 'unpredictable effects of local adoption'. Katz et al. noted that the effects of modifying a guideline to conform to local beliefs and local systems of care are often unpredictable and few studies address this topic. Therefore, it appears
from the literature that local modification or development is promoted to improve physician compliance with recommendations. At the same time, the effect on outcomes due to local practice guideline development is unknown.

OSUMC Practice Guidelines

This study examined the cost and effectiveness of drug therapy orders compliant versus not-compliant with two drug therapy practice guidelines developed and approved at The Ohio State University Medical Center (OSUMC). The two practice guidelines examined in this study included:

1. *Converting Intravenous to Oral Antimicrobials* – see Appendix A
2. *Criteria for Use of Vancomycin* – see Appendix B

The Appendices contain hard copies of the two guidelines, respectively. The two guidelines were selected based on the fact that each dealt with a significant drug therapy issue or problem. Also, each guideline also had an implicit goal. The first guideline, *Converting Intravenous to Oral Antimicrobials*, addressed the appropriate and inappropriate use of intravenous (IV) antimicrobial therapy. IV therapy is considerably more expensive than oral drug therapy. Thus, the implied goal of this guideline is cost reduction without impacting the quality of patient care. It has long been recognized that excessive and inappropriate use of antibiotics adds an unnecessary economic burden to the health care system. The second guideline, *Criteria for Use of Vancomycin*, also addressed the issue of appropriate use. In recent medical history, the overuse and
inappropriate use of antimicrobials has resulted in the development of bacterial
resistance.\textsuperscript{35,36} The implied goal of the guideline was the appropriate use of vancomycin
that should indirectly lead to a reduction in microbial resistance. Due to the extremely
low incidence of vancomycin resistance (less than 60 cases per year at OSUMC), this
variable was measured in this study but was not considered the main outcome.\textsuperscript{37} Rather,
the study focused on the appropriate use of vancomycin as defined in the practice
guideline.

\textbf{IMPACT OF PRACTICE GUIDELINES ON OUTCOMES}

Several general characterizations of the literature (for further discussion, see
Chapter 2) that were important for the purpose of this study were noted. First, practice
guidelines were rapidly becoming a tool used to state the standard of care for many
diagnostics, procedures and treatments. Second, there were several suggested uses for
guidelines. Yet, few studies have examined the cost effectiveness of guidelines. Third,
most guidelines were initially developed at a national level and modified to ‘fit’ local
clinical practice. The impact of local modification or development of practice
guidelines had not been studied. Fourth, compliance with guideline recommendations
was variable. Uncertainty in practice promotes the complying or not complying with
guideline recommendations. Finally, there were few studies of guidelines that were
conducted locally or that studied the cost effectiveness of complying or not complying
with recommendations.
No study had been conducted that examined approved guidelines at OSUMC. The fact that OSUMC guidelines were locally developed and implemented provided the need for study. The problem statement was particularly true for guidelines which focused on drug therapy issues. The intent of most drug therapy related practice guidelines has tended to focus on cost reduction and appropriate use of drugs. This intent was not completely without merit. Rather, the issue was that the effectiveness of care obtained by patients was not examined along with the drug therapy costs in many studies. Also, limited evidence existed in the literature linking compliance with guideline recommendations and cost effectiveness. Hence, the major purpose of the research was to quantify the cost effectiveness of drug therapy order compliance versus non-compliance with two locally developed and approved drug therapy practice guidelines.

RESEARCH QUESTIONS

What is the cost effectiveness of practice guidelines? The central research problem was independently addressed for the two guidelines — Converting Intravenous to Oral Antimicrobials (referred to as the iv to po guideline) and Criteria for the Use of Vancomycin (referred to as the vancomycin guideline) — by answering four research questions presented below. The study used decision analysis techniques to determine the cost effectiveness of drug therapy orders that comply versus not comply with the two guideline recommendations, respectively. Compliance was defined as a patient
treated by an order that followed the recommendation(s). The cohort of patients treated by orders that were compliant with the recommendations was referred to as the compliant group of patients (CGP). Non-compliance was defined as not being treated with a drug therapy order that followed guideline recommendation(s). Hence, the cohort of patients treated by drug therapy orders that did not follow recommendations was referred to as the non-compliant group of patients (NCGP). Each guideline was analyzed separately using the same set of research questions.

Each of the following research question sections provides information on the operational definition of variables and is be followed by specific research hypotheses used to answer the research questions. The study included the measurement of the variables listed in Appendix C. The variables were classified in the following six categories: patient demographics, clinical history, guideline compliance (or process variables), economic outcome, clinical outcome and health status.

Impact of Demographic Variables

The study collected information regarding the following demographic variables: age, gender, marital status, education level, payer type and admission service. The operational definitions for each variable are listed below.

age: age in years at the time of entry into the study

gender: male or female

marital status: married or single
education level: some high school, high school graduate, some college, college graduate, advanced college degree. NOTE: respondents were asked to record the highest education level obtained.

payer: Medicaid, Medicare, private managed care, private other, self pay, no pay. NOTE: primary and secondary payer(s) were recorded from the medical chart as it appears on the chart cover sheet.

admission service: surgical, medical, other. NOTE: The medical staff at OSUMC is organized by Divisions that are subcategorized into operational services. For the purposes of this study, the use of Divisions will be utilized due to sample size considerations.

The literature has shown that guideline compliance is affected by patient demographic variables. Hence, this study was designed to obtain information about selected demographic variables. If significant differences were found between CGP and NCGP on any demographic variable, the effect was either used for explanation of differences between the two groups or was statistically controlled in later analyses.

Impact of Clinical History

Ellrod et al. found that one of the main reasons for non-compliance with guideline recommendations was the severity of patient illness. Hence, this study collected information to check if patients in the NCGP were ‘sicker’ than those patients in the CGP or visa versa. If this was the case for any of the selected clinical history variables, then this variable was either used for explaining the differences between the two groups or was statistically controlled before conducting further analyses. The study collected data on the following clinical history variables:
comorbid conditions: presence of diabetes, renal disease, AIDS, active cancer, cancer history, organ transplant, cardiovascular disease or pulmonary disease. NOTE: The conditions listed expose patients to significant infection processes and were the same comorbid conditions used in a similar study examining the treatment of empiric infections.41

prior antibiotic usage: yes or no NOTE: Prior antibiotic therapy referred to therapy prior to study enrollment. Thus, it included 30 days prior to admission as the time frame.

days of present illness: number of days. NOTE: Days during previous 30 days patient was not able to perform normal daily activities

previous hospital admission: number of admissions in 30 days prior to admission in which enrollment occurred. NOTE: Current admission was included.

ED utilization: number of ED visits in 30 days prior to hospital admission. NOTE: Current ED visit was not included.

physician office visit utilization: number of physician office visits within prior 30 days to admission.

baseline health status: SF-12 item scores and two summary scale scores.

Thus, the demographic and clinical history analyses were conducted to identify differences between the two groups on these variables.

Guideline Compliance (Process Variables)

The study utilized the following variables to classify the patient into the CGP or the NCGP cohorts, respectively. The variables simply define which group a patient
belonged to as specified by the two respective practice guidelines (see Appendices A and B). The variables are presented for each guideline along with the designation of CGP or NCGP as defined in the guideline.

For IV Vancomycin:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>purpose of therapy:</td>
<td>prophylactic, empiric or laboratory confirmed treatment NOTE: If vancomycin was used for the prophylaxis of thoracic surgery, then the patient was classified CGP. If vancomycin was used for any other surgery prophylaxis then the patient was classified as NCGP.</td>
</tr>
<tr>
<td>endocarditis:</td>
<td>yes or no. NOTE: If yes and prophylactic then CGP.</td>
</tr>
<tr>
<td>thoracic surgery:</td>
<td>yes or no. NOTE: If yes and prophylactic then CGP.</td>
</tr>
<tr>
<td>beta lactam resistant gram + organisms (empirical):</td>
<td>empirical treatment yes or no. NOTE: If yes and subsequent lab tests confirm beta lactam resistant gram + and therapy continued then patient is CGP. If yes and subsequent lab test do not confirm beta lactam resistant gram + and IV vanco is discontinued then patient is CGP. If yes and subsequent lab tests do not confirm beta lactam resistant gram + and IV vancomycin continued then patient is NCGP.</td>
</tr>
<tr>
<td>beta lactam resistant gram + organisms (lab confirmed):</td>
<td>yes or no. NOTE: If yes, patient in CGP. If no, patient in NCGP.</td>
</tr>
<tr>
<td>beta lactam susceptible gram + organisms allergy:</td>
<td>yes or no. NOTE: If yes, patient is in CGP.</td>
</tr>
</tbody>
</table>
For IV to PO Switch:

- meningitis: yes or no. NOTE: If yes, patient is CGP
- endocarditis: yes or no. NOTE: If yes, patient is CGP

GI Function (Diet):
yes or no. NOTE: If patient was not taking an oral diet when the order was written, then CGP. If the patient was on an oral diet at the time the order was written then NCGP.

GI Function (Drug Therapy):
yes or no. NOTE: For patients yes to diet and drug therapy they were NCGP. For patients no to diet and drug therapy they were CGP. For patients with a mismatch of diet yes and drug therapy no or diet no and drug therapy yes additional data was gathered on the dates of first regular diet and first po medicine. If IV therapy continued after the date of a matched diet and drug therapy, then the patient was NCGP. If a mismatch continued throughout the course of treatment, then the patient was be CGP.

- date first regular diet: see above
- date of first po medicine: see above

Economic Outcome

For all cost variables used in the study, estimated cost was obtained. OSUMC cost estimates were calculated by using of cost-to-charge ratios. Currently, OSUMC does not utilize a cost accounting method. Therefore actual cost incorporating overhead costs directly into the measurement were not available. However, overhead costs were estimated at the operational level throughout the institution and have been incorporated into departmental cost-to-charge ratios. The term cost used in this study, thus, refers to
estimated cost. It is also felt that since the study was conducted at a single institution, cost estimates were appropriate since they were consistent between patients in the study.

Some of the variables listed below refer to resource utilization only. They are simple measures of resources consumed. For the purposes of the cost effectiveness analysis, only total estimated hospital costs were used for the cost portion of the ratio.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>total hospital cost:</td>
<td>total dollars. NOTE: Total hospital cost included all costs excluding physician cost associated with the hospital admission.</td>
</tr>
<tr>
<td>length of stay (LOS):</td>
<td>total number of days in the hospital during the admission in which enrollment occurred.</td>
</tr>
<tr>
<td>drug therapy cost:</td>
<td>total dollars. NOTE: Drug therapy costs were rolled into total cost and included only drug costs incurred in the hospital during the admission of enrollment.</td>
</tr>
<tr>
<td>laboratory cost:</td>
<td>total dollars. NOTE: Laboratory costs were rolled into total cost, and included only laboratory costs incurred in the hospital during the admission of enrollment.</td>
</tr>
<tr>
<td>days of lost activity:</td>
<td>number of days NOTE: Days during 30 days starting after the date of discharge that the patient was not able to perform normal daily activities</td>
</tr>
<tr>
<td>hospital readmissions:</td>
<td>numbers of readmissions. NOTE: Only readmissions for infection that occurred within 30 days after the date of discharge</td>
</tr>
<tr>
<td>ED visits:</td>
<td>numbers of ED visits. NOTE: Only ED visits for infection that occurred 30 days after starting the date of discharge</td>
</tr>
<tr>
<td>physician office visits:</td>
<td>numbers of physician office visits. NOTE: Only physician office visits for infection that occurred 30 days after the date of discharge.</td>
</tr>
</tbody>
</table>
new antimicrobial therapy: number of patients obtaining new antimicrobial therapy. 

NOTE: Only prescriptions obtained from a physician office visit 30 days after the date of discharge were included. No information was collected regarding the actual drug name, dosage or regimen due to patient recall concerns. The only data collected were if the patient had a new outpatient therapy or not.

Research Question One

Are economic outcomes for the CGP significantly different from the NCGP?

One of the main tenants of the use of practice guidelines for drug therapy products has been utilization rates (i.e., cost reduction). However, specific drug utilization may be a small part of overall costs incurred when treating a patient. Reduced drug utilization may produce short term cost efficiencies for an institution. This effect may delay cost incurred to the health care system. A patient may be re-admitted to a hospital, return for a physician office visit, return to an Emergency Department or start new outpatient antimicrobial therapy. In general, short term outcomes have been used to justify many guidelines. This fact has been the result of the institutional focus on short term economic outcomes. The economic variables used in the study addressed not only short-term institutional estimated costs but also longer term (30 day post discharge) resource utilization outside of the institution. Drug therapy order compliance and/or non-compliance with the guideline recommendations may have a significant economic
impact both inside and outside of the institution. The following specific research questions will be used to address research question one:

1.1 Do mean total hospital costs for the CGP differ from the mean of NCGP?
1.2 Do mean total hospital LOS for the CGP differ from the mean of the NCGP?
1.3 Do mean drug therapy costs for the CGP differ from the mean of the NCGP?
1.4 Do mean laboratory costs for the CGP differ from the mean of the NCGP?
1.5 Do mean number of days of lost activity in the first 30 days after discharge for the CPG differ from the mean of the NCGP?
1.6 Do mean number of hospital readmissions for infection in the first 30 days after discharge for the CGP differ from the mean of the NCGP?
1.7 Do mean number of ED visits for infection in the first 30 days after discharge for the CGP differ from the mean of the NCGP?
1.8 Do mean number of physician office visits for infection in the first 30 days after discharge for the CGP differ from the mean of the NCGP?
1.9 Is the proportion of patients receiving new antimicrobial therapy in the 30 days after discharge in the CGP different from the proportion in the NCGP?

Clinical Outcome

For the purposes of this study, two main clinical outcomes were examined - in-hospital mortality and antimicrobial therapy failure (success). The main effectiveness
measure used in the study is drug therapy success rate. The variable success was used in the denominator of the cost effectiveness analysis (CEA). For each guideline CEA, the success rate will represent the number of successful patients treated for infection (defined below) divided by the number of patients in the cohort group. For the iv to po guideline, failure was defined as the switch from iv to po therapy and then a subsequent switch back to an iv therapy or death. For the vancomycin guideline, failure was defined as the patient having continued symptomology of infection (fever, increased WBC and/or high pulse rate), the addition of another antibiotic within 3 days after initiation of therapy or death. Initiation of therapy was the same date as enrollment into the study. Karam et al. used similar methodology to define outcomes in a recent study of the use of vancomycin.\textsuperscript{36}

mortality: number of in-hospital deaths NOTE: The number of deaths per cohort group
drug therapy failure (success): number of failures(successes). NOTE: Although the number of failures per group was operationalized, the outcome of interest to the study was the number of successes. Hence, the compliment of failure is success. Failure is defined above.

Research Question Two

Do patients treated in the CGP have better clinical outcomes than the patients treated in the NCGP? Drug therapy order compliance with the guideline may produce clinically positive outcomes, negative outcomes, or have no effect. The main
effectiveness measure used in the study is antimicrobial therapy success rate. For each
guideline, the defined success rate will represent the number of successes divided by the
number of patients in the cohort group. For the iv to po guideline, failure is defined as
the switch from IV to PO therapy and then a subsequent switch back to an IV therapy or
death. For the vancomycin guideline, failure is defined as the patient having continued
symptomology of infection (fever, increased WBC and/or high pulse rate), the addition
of another antibiotic or death. The following specific research questions were examined
to address research question two:

2.1 Does the in-hospital mortality rate for the CGP differ from the rate in the
NCGP?

2.2 Does the drug therapy success rate for the CGP differ from the rate for the
NCGP?

Health Status Outcome

The SF-12 is a multipurpose short-form (SF) generic measure of health status. The
SF-12 measured eight health domains commonly represented in many health status
surveys and included: physical functioning, role limitations due to physical health
problems, bodily pain, general health, vitality (energy/fatigue), social functioning, role
limitations due to emotional problems, and mental health (psychological distress and
psychological well being). The SF-12 was developed due to concerns over respondent
The SF-12 generally takes 2 to 3 minutes to complete. The main purpose for the use of the instrument in this study was to capture a measure of health status and yet keep respondent burden low.

Although brief, the SF-12 accounts for more than 90% of variance as measured by the SF-36 instrument. The SF-36 is a longer instrument from which the SF-12 was developed. The scores for the SF-12 correlate highly with the scores for the SF-36. Ware et al. has compared the results of the SF-12 and the SF-36 in patients with four different medical conditions which differ from each other in terms of severity. For each of the four conditions, the SF-12 produced slightly less valid results, but the statistical conclusions for the two instruments were identical. Similar reliability and validity tests were conducted for patients with sixteen comorbid conditions, patients with five different acute conditions and various age groups of healthy people. As compared to the SF-36, the SF-12 produced the same statistical conclusions. The following two measures were obtained from the SF-12: health status and differences in health status. These two variables are defined as follows:

health status: SF-12 summary scales (Physical Component Summary (PCS-12) and Mental Component Summary (MCS-12))

difference in health status: difference between baseline and 30 day measurement for the SF-12 summary scores.
Research Question Three:

Do patients in the CGP report better health status outcomes than those in the NCGP? A significant feature of this study was the ability to examine the effect of order compliance with guideline recommendations on groups of patients' health status. All of the health status hypotheses were analyzed for the two summary scores of the SF-12 (i.e., the Physical Component Summary {PCS_12} and the Mental Component Summary{MCS-12}). The following specific research questions were addressed to examine research question three:

3.1 Are mean health status scores for the CGP different from the mean for NCGP at 30 days post discharge?

3.1.1 Does the mean SF-12 Physical Component Summary score for the CPG differ from the mean for NCGP?

3.1.2 Does mean SF-12 Mental Component Summary score for the CGP differ from the mean for NCGP?

3.2 Is the mean difference score (30 day post discharge score minus baseline score) for the CGP different from the mean difference score for the NCGP?

3.2.1 Is the mean difference SF-12 Physical Component Summary score for the CPG different from the mean for NCGP?

3.2.2 Is the mean difference SF-12 Mental Component Summary score for the CGP different from the mean for NCGP?
Cost Effectiveness

For the CEA part of this study, the variables used have already defined previously. Estimated total hospital costs using cost to charge ratios were used as the cost portion of the cost effectiveness ratio. The success rate for each guideline (as defined previously) was used for the effectiveness portion of the ratio.

Research Question Four:

The following specific research questions were addressed to examine research question four:

4.1 Is the cost effectiveness for the CGP different from the cost effectiveness for the NCGP for?

4.1.1 the iv to po guideline?

4.1.2 The vancomycin guideline?

ASSUMPTIONS

The study assumed that the clinical experts have done an effective job of reviewing current literature and developed the guidelines in accord with standard practice. However, the study will rate the two OSUMC guidelines on methodological standards established by the American Medical Association, Institutes of Medicine and
Canadian Medical Association. The rating and subsequent comparison of the OSUMC guidelines to a national study still will not judge the clinical quality or utility of the guideline recommendations.

The proposed research study also assumed that the implementation strategy used for the guideline was appropriate and effective. OSUMC utilized a published distribution process for implementation. Past studies have questioned the utility of providing physicians with printed materials and expecting behavior change. The study was bound by the process used at OSUMC. The study examined the cost effectiveness of drug therapy order compliance with guideline recommendations. These two stated assumptions above do not exclude the possibility of making development and/or implementation improvement recommendations based on the results of this study.

Another major issue for the study was the measurement of order compliance with the respective guideline recommendations. There was no attempt to measure this variable at the physician level. To do so was prohibited by sample size and power considerations. At OSUMC, there was an assumption that following the drug therapy guideline recommendations leads to positive outcomes. This assumption was the main premise of research for the study.
LIMITATIONS

Not included in the list of variables (see Appendix C) was a rating of physician’s knowledge regarding the guidelines. Although important, this variable was not central to the major research proposition posed by the study. The study was not designed to test implementation procedures. It was already assumed the guidelines were implemented satisfactorily.

A significant limitation of the study was the selection of the guidelines utilized for study. The results may be limited to hospital guidelines associated with antimicrobial drug therapy. The results may not be generalizable to other non-drug or drug class guidelines.

A second significant limitation of the study was the fact that the guidelines studied were institution specific. Although there was need to study locally developed guidelines, this aspect also limited generalizability. The results may be limited when making conclusions for other institutions. This limitation applied to the specific results reported but not the methodology of assessing the impact of locally developed practice guidelines.

The final significant limitation of the study and threat to external validity was the concern of selection bias. Since there was no randomization, one group or cohort of patients (compliant versus non-compliant) may be significantly different from the other. Although the study attempted to account for differences in socio-demographic and
clinical history, it was difficult to ensure the groups were equal in terms of disease severity. This concern was more apparent since this study used proxy measures of severity using recent clinical history variables.

DEFINITION OF TERMS

The definitions provided below were intended to serve as an aid to understanding key terms used in the study. The definitions were listed alphabetically. The key terms listed were for antimicrobial failure (success) rate, clinical practice guidelines, clinical practice guideline recommendations, cost effectiveness, decision analysis, health status, and order compliance.

Antimicrobial Failure Rate

For the purpose of this study, antimicrobial failure rates were defined for the two respective clinical practice guidelines. For the iv to po guideline, failure was defined as the switch from iv to po therapy and then back to an iv therapy or death. For the vancomycin guideline, failure was defined as the patient having continued signs and symptoms of infection, the addition of new antimicrobial therapy or death. The signs and symptoms measured in the study include fever, white-blood cell count and pulse. Thus, failure rate was defined as the number of failures divided by the number of patients in the cohort. The compliment of failure was success or the successful treatment of the infection.
Clinical Practice Guideline

Clinical practice guidelines were defined as, "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances". For the purpose of the study the terms practice guideline, clinical practice guideline and guideline were used interchangeably.

Clinical Practice Guideline Recommendation

Clinical practice guideline recommendations were defined in the study as the specific recommendations made by the two respective guidelines under study. The term recommendation will be used to refer to the aforementioned guideline recommendations.

Cost Effectiveness Analysis (CEA)

Cost effectiveness analysis (CEA) is one type of pharmacoeconomic evaluation that incorporates the resource input of an institution, program or treatment (i.e., cost) with outcome data (i.e., effectiveness). This study took the institutional perspective when measuring the cost and effect of order compliance or non-compliance with clinical practice guideline recommendations. This perspective was deemed relevant to the study due to the fact that the therapy covered by each guideline related to an acute event that occurred in the institution and was curable with proper treatment.
**Decision Analysis**

Decision analysis was defined as a method of cost analysis that uses a decision tree as a graphic tool to aid in defining cost-effective choices. The decision of writing orders to comply or not comply with the recommendations was studied as the two alternatives in the study.

**Health Status**

For the purposes of this study, health status was operationalized using the SF-12 health status questionnaire. The SF-12 health status questionnaire is designed as a generic tool to measure the patient's perceived health status at a single point in time. Studies on the reliability and validity of the SF-12 have been previously been published in the literature.\(^4\)\(^3\),\(^4\)\(^4\)

**Order Compliance**

Order compliance is a key term used to describe behavior with respect to following the practice guideline recommendations. Compliance was defined as following the recommendations. The cohort of patients treated with orders that were compliant with the recommendations will be referred to as the compliant group of patients (CGP). Non-compliance was defined as orders that did not following guideline recommendations. This cohort of patients was referred to as the non-compliant group of patients (NCGP).
SIGNIFICANCE OF THE STUDY

The study examined patient outcomes and cost effectiveness of order compliance with locally developed practice guideline recommendations. The historical assumption that a reduction in variation during the delivery of care has a positive impact on outcomes was tested by the study. The assumption that compliance with recommendations will lead to positive outcomes was also directly addressed. The results of this study may be useful when developing and implementing future practice guidelines. This will be especially true for guidelines which are designed to impact drug therapy decisions.

Given the large amount of resources utilized to develop practice guidelines, it was imperative for health care providers to know the value obtained. What was gained by the use of these valuable resources? If no effect in patient outcomes or cost effectiveness were found in the study, then complying or not complying with the guideline recommendations may be irrelevant. However, if effect were found, then behavior modification methods to increase compliance may be appropriate.

The study will generate new ideas for research. Although the study did not examine the impact of guideline implementation directly, the methodology used may be useful to evaluate different methods of implementation. The methodology may also be used to evaluate guidelines outside the area of pharmaceuticals. Finally the methodology may be used to evaluate two different guidelines which deal with the same clinical subject.
ORGANIZATION OF THE DISSERTATION

The rest of the dissertation was organized as follows. Chapter Two reviews the related literature on practice guidelines, the impact of guidelines on outcomes, health status and cost effectiveness. Chapter Three reviews research design, instrumentation, sampling, data collection and data analysis. Chapter Four presents the results that include a summary of the sample characteristics, psychometric properties of the health status instrument and the test statistics associated with each research question. Lastly, Chapter Five provides a discussion of the results obtained and reviews the implication of the results. Also, Chapter Five presents the limitations of this study and areas are suggested for future research.
REFERENCES FOR CHAPTER 1


37. The Ohio State University Medical Center. Unpublished Data. The Department of Pharmacy. 1999.


A great deal of information has been published regarding practice guidelines. Clinical practice guidelines have been widely produced and disseminated.\(^1\) There is intense interest, high expectations and limited information on effect of practice guidelines.\(^1,2\) At the same time, there is growing evidence of substantial unexplained and inappropriate variation in clinical practice patterns, growing concerns that further limitations in resources will reduce the possibilities to deliver high quality health care and growing concern that clinicians cannot assimilate rapidly evolving scientific evidence.\(^3,4\) In a survey of organizations who develop practice guidelines, the authors found the number one reason given for development of guidelines was to ‘improve the quality of medical care’.\(^5\) Within the literature, few cost or effectiveness evaluations have been completed. This Chapter begins with a review of the literature regarding the guideline process. In particular the processes of development, implementation and evaluation were reviewed. Included in the review of the guideline process, the main focus of the Chapter is on examining supporting literature for the research questions presented in Chapter 1.
Specifically, in the evaluation section, the following issues are examined with respect to their impact on practice guideline compliance or the impact of the practice guideline on outcomes: economic outcomes, clinical outcomes and health status outcomes and cost effectiveness. Finally, the Chapter ends with a general review of the health status and cost-effectiveness literature.

PRACTICE GUIDELINE PROCESS

The guideline process as defined by this study was presented in Figure 1.1. The process includes development, implementation and evaluation of the practice guidelines. This study focuses on guideline evaluation, but cannot be considered completely independent from guideline development and implementation. Thus, each of the three components of the process will be reviewed.

DEVELOPMENT

Millions of dollars in resources have been utilized on a national, state, and local level to develop practice guidelines. Practice guideline development is widely endorsed by professional health associations, government agencies, health insurance providers, hospital organizations, consumer groups and some health care practitioners. Thousands of practice guidelines have been developed. It has been estimated that 87% of all physician medical groups and independent practice associations are developing and implementing practice guidelines. The interest in practice guidelines
has been previously discussed. The purpose of practice guidelines is clear – to inform medical decision makers and to decrease variation in care by systematically influencing clinical decisions. In other words, change the way in which care is provided to patients. However, as the number of practice guidelines increase the danger exists to confuse practitioners. This confusion is created by different organizations with different methods and objectives developing guidelines on similar or even identical clinical topics.

The use of practice guidelines as an educational tool in an attempt to alter physicians' behavior is not a new concept in health care delivery. In the 1950's Lembcke et al. introduced the idea of developing explicit, objective criteria based on the medical literature to be used in medical care auditing. In the 1970's, Payne et al. stressed the importance of medical staff involvement in the development of explicit process criteria to improve adoption into clinical practice. Williamson et al. used a participatory model of criteria development involving not only physicians but other multidisciplinary health care personnel. Thus, the development of practice guidelines has a significant history. However, the amount of resources used to develop guidelines is relatively new and warrants examination.

There are numerous articles in the published literature which provide methodology standards to develop guidelines. Despite the voluminous literature on development, there is no single consensus method recommended for developing a practice guideline. Table 2.1 presents eight steps used to develop practice guidelines (Browman, 1995).
Steps to Develop Practice Guidelines

1. Select and frame the topic
2. Generate preliminary evidence based recommendations
3. Reconcile interpretations of evidence to ratify
4. Apply clinical modulating factors to formulate the guidelines
5. Have the guideline independently reviewed
6. Negotiate practice policies
7. Adopt practice guidelines and policy
8. Review and update the practice guideline

TABLE 2.1 Steps to Develop Practice Guidelines

**Topic Selection (Step 1)**

The first step is to select the topic for the development of a practice guideline. Very little aid is available in the literature on how to select a topic for guideline development. It is assumed that some criteria, either explicit or implicit, are used to set priorities. The criteria may include the burden of disease on population health, the state of scientific knowledge, the cost of treatment and the economic burden of the disease on...
A significant issue associated with the selection of the topic is the interest of practitioners or groups of practitioners in developing a topic. Without interest from experts, no topic would be developed into a practice guideline. Thus, one potential source for selection bias associated with guideline development is selection of the topic. Credibility of the final product, a practice guideline, may be questioned based on the choice of the topic. It is clear from the literature that topic selection is one area where research is needed to help guide the development process.

Evidence Gathering and Assimilation (Steps 2 and 3)

Steps 2 and 3 in Table 2.2 involve gathering information to formulate the guideline and ultimately make recommendations for clinical practice. Understanding the quality of scientific information that underlies each guideline recommendation enables a practitioner to judge whether the science supporting any recommendation is compelling enough to be adopted into practice. Woolf et al. presents four general approaches to practice guideline evidence assimilation:

A. Informal Consensus - a group of experts are convened and a subjective assessment of the literature is made and summarized into a guideline. This is the most commonly used method due to simplicity and flexibility.

B. Formal Consensus - a group of experts are convened and an objective assessment of the literature is incorporated into the guideline.

C. Evidence Based - criteria are established to include literature as ‘evidence’. A systematic process is used for locating and evaluating the evidence. The experts explicitly weight the evidence for all recommendations made in the guideline.

D. Explicit Approaches - the approach is similar to C, except no attempt is made to weight the evidence. Recommendations are made only when clear evidence exists in the literature to support it.
Methods A and B make no attempt to link literature evidence to recommendations made in guidelines. Hence, it is not surprising that the validity of guidelines developed using these two methods is often questioned. It has been said that consensus means that lots of people say collectively what nobody believes individually.\textsuperscript{3} It is no easy task to review large bodies of evidence and condense it into a few specific recommendations. Formal consensus development programs now exist in Canada, Britain, Sweden, Norway, Finland, Denmark, Holland, France and the United States.\textsuperscript{3} All of these programs use different methodologies to assimilate information.

Regardless of method used, it is important to recognize that the composition of the experts utilized in the development phase can profoundly influence the recommendations made.\textsuperscript{21} More importantly, the validity of a guideline is judged by ‘who’ developed it and what underlies their purpose. A major source of selection bias in the development process is who is gathered to do the developing. Practitioners must have faith in the author’s knowledge, experience and objectivity to prevent a lack of credibility.

Practice guidelines are ‘woven of judgment derived from clinical experience and are based on unequivocal scientific evidence at only a few points’.\textsuperscript{31} Evidence is often absent for many clinical decisions.\textsuperscript{32} Practice guidelines are an amalgam of clinical experience, expert opinion and research evidence. Lohr et al. Reports that the Institute of Medicine found that for approximately 4% of all health services the scientific evidence is strong, for 45% at best modest and for 51% very weak or non-existent.\textsuperscript{22} Lack of evidence has led to the emergence of a paradox in guideline development. As
more specific recommendations are sought in guideline development, the need for evidence increases. Where the evidence is lacking, a reliance on expert opinion also increases. The problem is the issue of validity. Biases introduced to a guideline in the development phase may reduce compliance or use by practicing physicians. Lohr et al. cautions that the state of scientific knowledge or empirical evidence about most ordinary health care today is not encouraging.22 Obviously, the more rigorous the literature and review methods used, the more evidence-based the practice guideline.30 Yet, wide gaps persist between published recommendations and actual practice.33

There is little guidance on the optimal way to mix expert opinion and literature based evidence.34 What is clear is that there must be a link to science in the literature. The difficulty of doing this may be one explanation of why sparse information exists in the literature on how to do this. Practice guidelines do not guide most parts of the clinical process nor do they address how the functions involved are to be performed by individual health workers, departments or teams.35 Thus, there is a certain degree of deliberate vagueness in all guidelines. Ford et al. notes that the development process leads to practice guidelines which often represent the lowest common denominator of acceptable care.16 To get support and endorsement from medical staff, controversial recommendations that represent a significant change from practice are left out or left ambiguous.16
Experts in given clinical specialties may not acknowledge a lack of agreement. A group of specialized experts may feel pressure to reach a consensus. It remains unknown whether clinically specific guidelines or vague ones produce better results. It is known that vague guidelines are unlikely to influence provider behavior. Rice et al. suggested that guidelines based on poor evidence may have less effect on patient care. It remains unclear in the literature if consensus based guidelines are more or less effective than evidence based guidelines.

Practice guideline development is a result of reviewing all available scientific 'evidence' for a topic under investigation. Expert opinion is used to fill in gaps in the evidence. One unintended benefit of guideline development may be the recognition of clinical areas where gaps in the clinical literature exist.

Research comparing the effectiveness of different group judgment processes on how best to put medical evidence into guidelines is not commensurate with the effort to produce the guideline. There is very little evidence of which method is the most productive. Shekelle et al. conducted a study that compared the recommendations which resulted from using two different methods of guideline development. The authors compared informal consensus versus an explicit approach. The experts utilized did not differ on satisfaction with method used. However, the explicit method produced more specific guideline recommendations than informal consensus. Thus, the need for evidence increases as specificity of the recommendations is desired.
How and by whom the evidence is collected and collated into recommendations are significant issues in the development process. The importance of an underlying link to the scientific literature has been expressed by many authors in the literature.\textsuperscript{1,32,39}

Population Definition (Step 4)

Step 4 involves defining the population which the guideline is intended. There is no information in the literature to review concerning this issue. Several authors make the point that the population must be well defined.\textsuperscript{3,4,27} It is clear that not only the population to which the guideline applies but also the population to which the guideline does not apply must be well defined. It is not inconceivable that a practice guideline recommendation could be applied to a “wrong” patient.\textsuperscript{32} This topic is another area which needs more research to determine the effects of misapplied practice guidelines.

External Review (Step 5)

There is little information in the literature regarding the external review process to be used in the development process. The literature is clear that external review should be done.\textsuperscript{4,8,27} Who and how many should do external review has not been established. Often specialty groups develop practice guidelines but the intended audience or users of the guidelines are general practitioners. This issue has raised the issue of who should review what. Should the content or the practicality of the guideline be reviewed? The answer is probably both and practice guidelines should be reviewed by multidisciplinary practitioners.
Adoption of the Practice Guideline (Steps 6 and 7)

Steps 6 and 7 will be covered in the implementation section of this review. The point is that no matter how well developed a practice guideline is, if the decision tool is not implemented and adopted in practice, there will be no impact on practitioner behavior or on patient outcome.

Review and Update of the Practice Guideline (Step 8)

It is not surprising that as current knowledge expands, guidelines must be kept current and done so in an efficient manner. The literature notes that when new information is available which contradicts a guideline recommendation, then the update must occur immediately.\(^3\) At a minimum, practice guidelines and the literature should be reviewed on an annual basis.

National Versus Local Development

The national guideline developed by Middleton, et al. provides a perspective on developing a national practice guideline.\(^40\) The guideline developed was titled, ‘National Cancer Guideline for Prostate Cancer’. A major limitation to a national process is the need to find consensus. However, it should not be expected to get a consensus recommendation at a national level. The Prostate Cancer Guideline makes no specific recommendations. Middleton concludes that “practitioners should be aware of all (treatment) options”.\(^40\) Many nationally developed guidelines are vague in both
language and recommendations. National practice guidelines are also limited in effect
due to a lack of familiarity and trust in the developers.

The response of many local organizations is to modify national guidelines or
develop in-house or local guidelines to make them more accommodating to local
clinical practice. There is evidence that clinicians are more likely to follow practice
guideline recommendations if they were involved in the development process. Local practice guidelines may be more successful in changing behavior by facilitating
acceptance and compliance. Local modification also allows the local organization
to provide more specific recommendations. The paradox of the issue lies in the fact that
on the one hand, the hallmark of valid guidelines is the link to the underlying science.
Local modification or development may break the link to the science. On the other
hand, unless national guideline developers understand practical local issues, their work
may be ignored and doomed to failure. Brown et al. locally modified and adopted the
national AHCPR Depression guideline. The result was a document which was smaller
and more user-friendly. The national practice guideline was 327 pages while the local
version was 44 pages in length. The national guideline focused on background and
supporting documentation. The local version focused on practical information such as
screening information, drug therapy information and specific recommendations. One
advantage to modification was that the local version was not created from scratch.
Rather the credibility of AHCPR was retained. Despite the stated benefits, the
effectiveness of the modified guideline was not reported. The authors note that
although they had developed a more user-oriented document, there is no way to know if
any new bias, omission or error was introduced by the local process.\textsuperscript{32} The effects of modifying a national guideline to conform to local beliefs and local systems of care are often unpredictable and few studies address this topic.\textsuperscript{47,48} When a practice guideline is developed or adopted locally, the guideline may be different in content and quality.\textsuperscript{5,49-51} Therefore, it appears that local modification is promoted to improve physician compliance with recommendations. At the same time, the effects on compliance with and on patient outcomes remain unknown.

**Barriers To Development**

Grilli et al. identified the following list of problems associated with the development of practice guidelines.\textsuperscript{47} Practice guidelines:

1. tend to be excessively complex and difficult to test in practice
2. have limited applicability
3. fail to capture subtle clinical nuances that make a guideline recommendation inappropriate
4. tend to use poor quality scientific evidence
5. have unexpected effects on other aspects of care
6. fail to identify organizational inefficiencies that impede practice guideline implementation

The authors note these problems should be recognized and solutions incorporated into the guideline development phase.
Practice Guideline Development Quality

As presented in Table 1.1 (see Chapter 1), standards have been set for the development of practice guidelines. Based on these standards, Shaneyfelt et al. evaluated 279 peer-reviewed, published practice guidelines. Using yes or no, dichotomous responses to establish adherence to the twenty-five standards, overall average adherence was 43.1%. In other words, the average published guideline met 10.8 out of 25 standards. Moreover, adherence did not improve over time nor was adherence any different for guidelines produced by specialty medical societies, general medical societies or government agencies. The authors did find that longer written guidelines positively correlated to adherence to more methodological standards. The authors concluded that "published, peer-reviewed guidelines do not adhere well to established methodological standards." Further, the greatest opportunity for improvement is needed in the identification, evaluation, and synthesis of scientific evidence. Less than 10 percent of the guidelines reviewed used or described formal methods of combining scientific evidence or expert opinion. Twenty-five percent did not cite any references. Shaneyfelt et al. notes the instrument does not weigh the standards differentially and the yes/no format does not assess relative standard compliance with a given standard. However, the instrument was rigorously tested for reliability and validity. For the purpose of this study, the instrument developed by Shaneyfelt et al. was utilized to compare the OSUMC guideline development process resultsto those published in the peer-reviewed literature. The methodology of this process will be discussed later in the paper (see Chapter Three).
IMPLEMENTATION (EDUCATION)

Once a clinical practice guideline has been developed, implementation of the guideline must occur to meet the guidelines stated purpose and goals. Before a practice guideline can improve quality or outcome, reduce inappropriate utilization (both over utilization and underutilization), control escalating expenditures, reduce practice variation and facilitate clinical application of growing scientific knowledge, the guideline must be implemented and used. Practice guidelines can contribute to improved care only if they succeed in moving actual practice closer to the behaviors the guideline recommends. The attention paid to implementation to date has been inadequate to date. Implementation includes (but is not limited to) the effective education of practitioners. The effectiveness or impact of a guideline rests in part on the ability to implement it. The most evidence-based and specific guideline recommendations will not produce results if poorly implemented. Conversely, the least evidence-based recommendation may be followed rigidly with proper implementation. Practice guidelines may do more harm than good if they are inappropriately implemented or applied.

In a survey of organizations developing guidelines the authors found that most organizations did not actively pursue implementation. Kibbe and colleagues have described what they term the ‘insufficient dissemination paradigm’. This paradigm posits that little behavior change by practitioners occurs or should be expected to occur
by simply making available printed information. Studies have shown that simple
distribution of a written document does not change clinician behavior.\textsuperscript{3,32,52,57,58}
O'Malley et al. stated that 'physicians do not pay attention to protocols unless they are
linked to financial incentives'.\textsuperscript{59} Further, the author states that 'when we send out
guidelines - and that's all they are, guidelines - they travel at the speed of sound from
the hand to the wastebasket'.\textsuperscript{59}

Methods of Implementation

There are several methods used to implement practice guidelines. These
methods include dissemination of educational material via print, audio or video media
and by live presentation. Dissemination is the most popular method of implementation
due to low cost and convenience. However, as previously noted, dissemination may not
have the desired effect on behavior change. In fact, Eddy et al. states that it is purely an
assumption that synthesizing and conveying the available scientific knowledge as a
message in a practice guideline to clinicians through a credible source and in an easily
understood form might improve provider performance.\textsuperscript{51} It can not be stated enough
this assumption has not been supported by current literature. Davis et al. reports that the
creation of guidelines without attention to their adoption is a sterile exercise which
wastes precious intellectual, human and financial resources.\textsuperscript{60} The mere dissemination of
guidelines may predispose practitioners toward change, but is not enough to alter actual
practice.\textsuperscript{52} Other methods of implementation include use of local opinion leaders,
utilization of computer technologies mainly in the form of reminder notices, counter
detailing by health professionals and using local consensus processes. Many of these methods will be discussed further in the Evaluation Section below. Implementation of practice guidelines is difficult, and the ideal method has not been identified.\(^{35,53,58}\)

**Medical Information Diffusion Models**

The literature has studied the diffusion of new information models into the medical community. The purpose of this review is not to detail medical information diffusion. That is not a purpose of this study. Rather, a brief review will aid in the discussion of the Implementation Studies Section. Traditional diffusion models rely on a rational, information-seeking and probabilistic health care practitioner to go get new information. These models hold that dissemination works based on the presumption that synthesizing and conveying scientific evidence through a credible source can improve clinical practice.\(^{47}\) The model views the health care practitioner as an active consumer of new information. It is naturally a person who wants to intrinsically keep up with new information. The practitioner is viewed to devote time and effort to the task. When new information is found, they are willing to change how they practice. Unfortunately, there is considerable evidence that this scenario is wrong.\(^{61,62}\) Too much faith has been placed on the mere activity of producing new information and giving it to the receiver of new information.\(^3\) It was stated earlier that simple distribution of material does not have much impact on behavior.\(^{32,39,58}\) A recent study compared the result of implementing a practice guideline by two methods on pharmacy costs, health care utilization and blood pressure.
control. The first method consisted of dissemination of the practice guideline. The second method included dissemination, use of a clinical champion, special education procedures, participation of a clinical pharmacist and feedback. The results found no difference in pharmacy costs or health care utilization. The study did find an increase in the use of guideline recommended medications and a decrease in non-guideline recommended medications.

Newer models of information diffusion involve an examination of a complex process of information exchange. These models study interactions between characteristics of receivers, the service, the message and the channels of information. However, as will be seen in the studies presented below, successful diffusion does not automatically result in practitioner behavior change.

Davis et al. conducted a literature review to recommend effective strategies for implementing practice guidelines. The authors found the following factors affect adoption: quality of the guideline, characteristics of the health care professional, characteristics of the practice setting, incentives, regulation and patient factors. The authors conclude that weak effects are obtained from didactic, traditional continuing education and mailing dissemination; moderate effect from audit and feedback (concurrent), targeted to specific providers, delivered by peers or opinion leaders; and, relatively strong effect from reminder systems, academic detailing and multiple interventions. Serious deficiencies exist in the adoption of practice guidelines in practice. Weingarten et al. observed when active implementation strategies are stopped, practice behavior reverts to that observed before implementation of an intervention.
Studies of Implementation (Compliance)

The most common method used to test implementation strategies is to examine compliance with recommendations of practice guidelines before and after a method has been applied. To aid in compliance and implementation, Kelly et al. suggests that guidelines should be developed at a local level, use intensive educational strategies, provide the clinicians with feedback and use local opinion leaders. Research indicates some of these strategies work some of the time.

The main assumption of guideline development is that adherence or compliance with a practice guideline will improve the efficacy and effectiveness of care by reducing variation. In fact, the use of practice guidelines as standards raises the question of compliance. Should compliance be 80%, 90% or 100%. Does compliance depend upon the severity of the consequences if not followed? Some authors even support the notion that compliance may in and of itself be a quality indicator. Thus, compliance with guideline recommendations is a significant issue to the guideline process. Physicians may or may not comply with practice guideline recommendation(s) for both appropriate and inappropriate reasons. Thus, compliance will not ever be 100%. Absolute compliance with practice guidelines implies that physicians may blindly follow recommendations that may be a threat to patient outcomes and quality of care.

In a review by Grilli et al. twenty-three studies were included in which practice guideline compliance rates were reported for guidelines endorsed by official organizations. A total of 143 recommendations were made by the published guidelines. The authors found that the mean compliance rate with the 143
recommendations was 54.5%. The subspecialty groups of cardiology and oncology had the highest rates of compliance. A significant result was that as the complexity of the guideline increased compliance decreased. This review was supported by other research in the diffusion of new information literature which has shown that certain characteristics of a recommendation may influence compliance.

Research to determine the impact of practice guidelines on professional behavior has yielded disappointing results. Even when there has been documented evidence of increased physician knowledge, the change in behavior has been difficult to demonstrate. For example, physicians who received (via United States mail) continuing education on hypertension did not show an improvement in lowering the blood pressure of referred patients. This result was obtained despite a documented increase in knowledge. Thus, there appears to be no lasting influence of short-term knowledge gain. In another study, when information on the effectiveness of drugs for congestive heart failure (CHF) were supplied to selected physicians, compliance with the guideline was not higher than physicians without the information. In the study, physicians were randomly assigned to a control group or an intervention group. Physicians in the intervention group were mailed a cover letter and questionnaire, journal article and the history profile of a CHF patient who would benefit from the information. The results indicated that two-thirds of the physicians were aware of the published information and one-third intended to change their practice based on the information. However, there was no difference in actual behavior between the control group and the intervention group. The authors conclude that physicians who receive educational interventions after
the publication of a study were no more innovative in prescribing medicine for CHF patients than physicians who did not. They found that the strongest predictor of compliance was the physicians subjective a-priori intention to comply.

In a before and after study of the dissemination of printed practice guidelines to reduce the use of cesarean sections, Lomas et al. observed most obstetricians were aware of the new guideline recommendations (94%). Most agreed with the guideline recommendation (85%). Thus the attitudes of the obstetricians in the study were congruent with the guideline before and after dissemination. After distribution, one-third of the obstetricians and hospitals self-reported changing practice as a result of the guideline. Physicians reported rates of cesarean sections in women with previous cesarean sections that were significantly reduced as compared to prior to dissemination (from 72% to 61%). This was as the guideline recommended. However, actual knowledge of the guideline recommendation was poor (67% gave correct responses to scenarios involving the guideline). The problem is that real data on actual practice showed rates of cesarean section that were 15 to 49% higher after dissemination. The authors conclude that although the practice guideline may predispose physicians to consider changing behavior, unless other incentives or removal of disincentives were employed there was no change in actual behavior. In a separate but similar study, the
author found similar results in that physicians are apt to report high compliance with practice guidelines but when actual practice is examined, few actually follow the recommendations.\textsuperscript{76}

In a before and after study of dissemination of the 1984 Report of the Third Joint National Committee on the Detection, Evaluation, and Treatment of High Blood Pressure, the authors found that 62\% of respondents were aware and 81\% were knowledgeable of the recommendation prior to dissemination.\textsuperscript{77} Thus, only 17\% reported using the guideline in their practice. No differences were noted after dissemination since most participants in the study were compliant with the recommendations prior to guideline development. The authors concluded that the guideline codified behavior already occurring and no change was possible.\textsuperscript{77}

Gorton et al. evaluated the effect of four methods of dissemination of asthma practice guidelines on physician behavior and attitudes toward the dissemination strategies.\textsuperscript{53} The treatment groups consisted of physicians who were 1. Given a short summary phone call by detailing physicians and a continuing medical education (CME) conference, 2. Computer reminders and CME conferences, 3. A multimedia approach (facsimiles, posters, video, audio and CME conferences), and 4. CME conferences only. The results indicated that all groups improved in the use of medicines or improved the use of peak flow meters. However, none of the methods resulted in improvement for both process variables. The authors concluded that dissemination should be done with multimedia emphasizing short, concise summaries with frequent reminders.\textsuperscript{53} The
authors also note the inability of practice guidelines to be translated into practice could undermine the role of practice guidelines as a strategy for quality improvement, cost containment and health system reform.\textsuperscript{53}

Even active participation (i.e., as an expert) in the development of practice guidelines under the National Cancer Institute (NCI) funded Community Hospitals Oncology Program did not increase physicians adherence to appropriate staging, medical management, nursing and rehabilitation behavior for the treatment of breast, rectal, and lung cancer.\textsuperscript{16} The developed practice guideline was disseminated in paper form by attaching it to the pathology report of newly diagnosed patients which were then placed on the patients' charts. This study tests the assumption that local community physician participation in guideline development and dissemination would enhance the adoption of the guideline. The results failed to support this assumption.\textsuperscript{16} The authors concluded that local practice guideline processes alone had minimal effect on behavior.\textsuperscript{16} The authors state, "practice guideline development appears to be a time consuming and relatively ineffective approach to diffusion of state-of-the-art management into practice. Loosely structured organizations to set up consensus practice guidelines are probably a poor educational investment"\textsuperscript{16}. The change in behavior (compliance) appears to occur in a slow and incremental fashion. However, a recurrent theme is that practice guidelines issued by national and state organizations are unlikely to have much effect unless they are translated into local or individual action.\textsuperscript{61,57,58} There is need to follow-up after dissemination to help translate the message into change behavior.\textsuperscript{79}
Pathman et al. conducted a study to examine the cognitive and behavioral steps taken to comply with a national guideline with pediatric vaccinations. The authors used the awareness to adherence model. The model posits that for adherence to occur, the physician must first become aware of the guideline, then agree with it, and then finally adopt it into clinical practice. The study found general support for the model. Physician characteristics and beliefs have a significant impact on compliance with guideline recommendations. The study questions the idea that non-compliant physicians are uninformed, uninterested, entrenched in old habits or unwilling to act. Rather, physicians simply do not agree with the guideline recommendations. It is too simplistic to think that giving physicians information will change behavior.

In a survey to identify characteristics and attitudes which led to guideline compliance it was found that physician knowledge did not predict compliance behavior. The main predictor of compliance was the physician's belief that they could change patient behavior through counseling during a patient visit. Similar results were found in a study by Cohen et al. where the single most important predictor of behavior after the intervention was the physician's behavior prior to the intervention. This study again illustrates that increased knowledge had no impact on behavior.

In a retrospective case analysis to assess the compliance of physicians with published local guidelines for the diagnosis of interminute lung scans and for the treatment of pulmonary embolism it was found few clinicians followed the recommendations (compliance rate less than 12% for the diagnosis guideline and 53% for the treatment guideline). The authors conclude that underutilization of objective
imaging methods for establishing the diagnosis implies over-reliance on clinical impressions which are known to be unreliable.

Carpenter et al. examined the determinants of compliance with practice guideline recommendations. After a guideline in the appropriate and inappropriate use of glucocyte colony-stimulating factor was disseminated, the authors found that age can influence compliance. The study also found that physicians who were most knowledgeable about the guideline, override the recommendations the most. The authors posit that these physicians may be the most comfortable to do this.

In a study comparing the National Asthma Education Program (NAEP) guideline recommendations to actual care received by a large independent practice association (IPA) type HMO population, patient compliance with the guideline was poor. Patients treated by an asthma specialist tended to be more compliant. The study found compliance varies by patient characteristics (age, race, disease severity, etc.). However, the study did not determine if the differences were due to patient characteristics per se or to differences in care provided to groups of patients within demographic groups.

**Barriers To Implementation (Compliance)**

It has been noted that frequently physicians do not follow recommended standards of care. Several studies report that practice guideline compliance is incomplete and inconsistent. The studies reviewed above demonstrate inconsistent results for different methodologies.
The lack of commitment and support by physicians for practice guidelines occurs for three valid reasons:91

A. lack of outcomes studies
B. limited availability of information systems, and
C. time and expense involved in development

Other barriers to compliance in practice exist, including poor planning, poor allocation of resources, over-emphasis on practice guideline development and failure to attend to guideline implementation. Thus a major barrier to implementation is the cost and complexity associated with it. Since most of the resources are devoted to development, there is little left for implementation.

A barrier for many organizations is the fact that the medical community is uncertain whether practice guidelines are necessary or even helpful.7,92-94 In other words, the validity of the guideline documents is often in question. Some feel that traditional medical practice and attitudes are resistant to the concept of practice guidelines. Different organizations have legitimately different interests and priorities that create difficulties when implementing guidelines.

Gifford et al. suggested that clinicians resist practice guidelines seen as a threat to autonomy.95 There is a distinction between factual claims (or what is) and normative claims (or what ought to be). Practice guidelines make normative or prescriptive claims, while outcomes data is factual. To infer a normative claim (practice guideline) from outcomes data (fact) is to implicitly have a normative assumption (i.e., the goal of the guideline is appropriate). There may be valid concerns by providers that payers and
courts may use guidelines as evaluation tools for short term economic and quality reasons, respectively. In fact, tort liability may increase if defective practice guidelines are followed blindly and cause bad patient outcomes.\textsuperscript{28}

Not all guideline recommendations will be followed. This ‘non-compliance’ is often appropriate.\textsuperscript{96} In fact, most practice guidelines should only apply to 60 to 95 percent of relevant cases.\textsuperscript{97} The key is that the guideline must be developed including criteria where the guideline does and does not apply.

In a study which retrospectively examined a group of physicians defined as non-compliant with length of stay guidelines, the authors found that physician refusal to follow guideline recommendations account for a small percentage of non-compliance (16 percent).\textsuperscript{67} The practice guideline studied was a length of stay guideline for low-risk patients admitted with chest pain. Physicians were given risk information concurrently by letters placed on the chart and physician-to-physician telephone calls. The reasons accounting for the majority of non-compliance include implementation issues (42 percent), health system inefficiencies (14 percent) and changes in severity of patient illness (9 percent). Implementation issues involved patients who were misclassified as low risk (both false positive and false negative). Health care system inefficiencies included waiting for the return of diagnostic tests, procedure delays, social issues, consultation delays and patient refusal for care. Thus, physician refusal accounted for a small percentage of the non-compliance. Specific clinical or logistic reasons exist for the majority of non-compliance with the guideline.
Non-compliance with practice guideline recommendations occurs for a variety of reasons. Some of the reasons are valid, others may not be valid. The key issue is that different clinicians use different criteria for choosing to comply or not comply with recommendations. Greco et al. found that when cost reduction is the goal of a guideline, physicians respond more favorably to evidence that the proposed change’s improve patient clinical outcomes or at a minimum patient outcomes are not worsened.\textsuperscript{44} It is impossible for any practice guideline to delineate all of the clinical subtleties of a real patient.\textsuperscript{57} Litzelman et al. found when they assessed the reasons for non-compliance with computer generated preventive care reminders, 22.6\% of the physicians felt it was not applicable (e.g., the test was done elsewhere or the patient was too ill), 27.5\% said they planned the test for the next patient visit (e.g., physician was too busy during this visit, patient requested it for the next visit), and 9.9\% the patient refused the test (e.g., felt the test was unnecessary, too costly, patient was too busy, patient did not want to know the results).\textsuperscript{96} The authors found that by requiring physicians to explain non-compliance, compliance gradually increased 19\%.\textsuperscript{96} 

The conclusion drawn from the implementation literature is that no single method(s) of implementation have been successful in changing physician behavior. Dissemination must occur but how it should occur has not been determined.

\textbf{EVALUATION}

Despite the enthusiasm and opportunities for quality improvement based on guidelines, far more effort and resources have been applied to development of
guidelines versus the evaluation and implementation of guidelines. There are few examples in the literature of how practice guidelines improve the quality of care or reduce health care cost. Expectations about practice guidelines impact on care may vary depending upon perspective. Clinicians, patients, payers, administrators and politicians all have very different expectations regarding what practice guidelines should do. The problem is that evidence is lacking that confirms that practice guidelines are meeting any perspective's expectations. All perspectives want improved quality of care from guidelines. Yet, patients are concerned with clinical outcomes; physicians want improved outcomes but also want limits on intrusion on autonomy, income and liability; insurers, employers and public agencies want reduced cost and variation; administrators want reduced cost, increased efficiency and improved risk management; politicians want reduced public expenditures; utilization reviewers want to identify inappropriate care; and, attorney's want help proving negligence. No guideline can satisfy all of these expectations. Carter et al. found that when organizations who develop practice guidelines were asked, less than 13% of the organizations reported doing any form of evaluation. Therefore, the utility of practice guidelines remains a mystery and often fail to achieve stated objective(s). Guidelines will only be useful if they are utilized as part of comprehensive programs to improve quality. To date, practice guidelines have not played a major role in cost containment or quality improvement efforts. Just as the previous section discussed implementation barriers this section will report evaluation barriers. The combination of these barriers results in a lack of use for quality improvement.
Johnston et al. provides an overview of how to study the effectiveness of practice guidelines. Two perspectives are possible. The first is a process-focused approach where a researcher focuses on variables assumed or known to lead to the goal of the guideline. The second approach is outcomes focused. The latter approach studies the final outcome variables that are goals of the guideline. Guidelines should specify key measures of outcomes or process variables to be monitored. However, little is currently known about the real effects of guidelines versus other confounding effects. Doubt exists about the ability of a practice guideline to reduce costs or increase quality. In fact, strategies to scientifically evaluate the effectiveness of practice guidelines in clinical settings have largely been ignored.

Given this fact, a brief selection of studies will be reviewed.

Research on the development and implementation of guidelines has been previously covered. This section will focus on published studies of effect of outcome. Although the Implementation Section contains brief discussion of attitude, knowledge and behavior, these topics will be discussed in detail here. The former discussion focused on the methods of implementation. The current discussion focuses on the effects. Figure 2.1 depicts a simplified conceptual model of the discussion in this section.
A brief review of practice guidelines impact on knowledge, attitudes and behavior will be presented. The focus will be on outcomes. Woolf et al. presents the potential benefits of each of the parts listed in Figure 2.1. Practice guidelines may benefit knowledge by improving medical education and defining needed areas of research; benefit attitudes by promoting rapid acceptance of new standards of care and
enhance credibility; benefit behavior by increasing practice patterns with recommended practice and reduced variation; and, benefit outcomes by improving clinical outcome and reducing cost (e.g., improved value). The potential exists, but little evidence exists for the latter outcome benefit.

Methodology of Evaluation

It has been argued in the literature that evaluation of practice guidelines should be conducted by rigorous methodology similar to clinical studies. However, other authors note that due to difficulty in manipulating and controlling an entire stock of confounders, traditional clinical investigation strategies are difficult to apply. It may not be possible to evaluate practice guidelines in a generalizable manner. Program evaluation is one common strategy used to evaluate practice guidelines that are embedded in a naturally occurring setting. This type of evaluation occurs after the implementation of the guideline. The goal is to evaluate if the practice guideline is meeting stated objectives. The evaluation may focus on process, outcomes or efficiency. This type of evaluation is the main type of evaluation that appears in the literature.

Attitudes

There appears to be very little evidence that practice guidelines impact attitudes of practitioners. Stange (1992) et al. found in a survey that physicians disagreed with 12% of the published guidelines presented to them. In this mailed survey the
authors measured physician agreement with the recommendations of the U.S. Preventive Services Task Force (USPSTF) guidelines. The average physician agreed with 88% of the recommendations. Disagreement was associated with older age, not having completed a residency, male sex, less prior exposure to the USPSTF guidelines and greater perception of the impractibility of applying them. Thus, provider characteristics may influence the attitudes towards a practice guideline. The Hill et al and the Lomas et al. studies presented previously in which the clinicians already held attitudes congruent with the recommendations made by each guideline, respectively.\textsuperscript{3,77}

Wilson et al. examined attitudinal barriers that may exist which prevent the adoption of practice guidelines.\textsuperscript{105} In a mailed survey, 23% of the respondents reported using guidelines. Sixty-six percent had confidence in a practice guideline if developed by a physician group. Only 5% had confidence in federal agency based or third party payer developed guidelines. Older physicians also ascribed more negative qualities to practice guidelines than younger physicians. Physician attitudes towards practice guidelines are associated with the identity of the issuing organization and the age of the practitioner.\textsuperscript{105}

Hayward et al. conducted a survey that examined physician attitudes towards guidelines.\textsuperscript{108} The study found general positive attitudes with more than 50% of respondents reporting guidelines to be a convenient source of information and good educational tools. However, 22% were concerned about loss of autonomy. Fifty-one to 77% were not confident in guidelines issued by federal or provincial health ministries or by health insurance plans. The study concludes that although physicians feel positive
about guidelines, they have not integrated their use into practice. In a similar U.S. study by Tunis et al. respondents reported guidelines were good education tools (64%), convenient sources of evidence (67%) and developed to improve quality of care (70%).

The respondents also report guidelines were 'cookbook medicine' (25%), too rigid to apply to real patients (24%), were a challenge to autonomy (21%) and would be used in disciplinary actions against physicians (68%). This study found mixed attitudes towards guidelines. An interesting note was the fact that the study found more positive attitudes to be held by physicians in non-private practice, who were salary based and who saw patients less than 20 hours per week. Ferrier et al. found similar results in a Canadian study.\(^{109}\)

Thus, most clinicians already have the attitudes espoused by many practice guidelines prior to development. It is no surprise that few studies have shown a change in attitude post implementation.

**Knowledge**

It has been shown that clinicians' knowledge increases with distribution of practice guidelines.\(^{77,110}\) Stange et al. found that within one year after the dissemination of a practice guideline, 60% of clinicians were aware of it.\(^{107}\) Awareness can be improved through the use of aggressive dissemination strategies. Results of the Lomas et al. study have already been presented.\(^3\) These authors found that 2 years after the dissemination of a cesarean section guideline, 87% of obstetricians were aware.
In a British study, where a practice guideline on smoking cessation counseling was mailed to all general practitioners, the authors found that 51% remembered getting the guideline. Twenty-eight percent read it but only 9% could name one of the three recommendations present in the guideline. In a similar study, Sadowsky et al. noted that when dentists were mailed a laminated placard with the American Heart Association practice guideline on prophylaxis antibiotics, knowledge increased significantly. Those who got the mailed placard had a significantly higher knowledge performance on 5 of 6 vignette cases as compared to the dentists who did not receive the placard. In a similar U.S. study, familiarity with practice guidelines ranged from 11 to 59% depending on the specific guideline. Subspecialty physicians were more familiar with practice guidelines associated with their own subspecialty. General internists were more familiar with general diagnostic guidelines.

Knowledge appears to be increased by practice guidelines. However, the level of knowledge gained appears to be variable. The results have shown thus far that attitudes appear to be positive with respect to guidelines and knowledge increases with guidelines. The next section examines if these two trends translate into behavior change.

Behavior

It has been stated several times in this report that clinicians often fail to follow recommendations. There are six basic methods used to change physician behavior:
1. education
2. feedback
3. participation by physicians in efforts to bring about change
4. administrative rules
5. financial incentives
6. financial penalties

The first four methods have been tried and tested with respect to guideline implementation. Greco et al. noted that no single method is inherently effective, especially when used alone. Strategies that rely on more than one method appear to be more successful. It remains to be determined if any of the methods change behavior or impact outcomes.

Behavior – Observational Studies

McPhee et al. found that despite well-publicized and implemented practice guidelines on screening mammography for the prevention of breast cancer, only 13% of the patients meeting criteria actually had a mammography performed. A similar study, the National Cancer Institute found that between 14 to 42% of women eligible did not have a screening mammography. This study found that the number one reason (for more than 85% of the respondents) reported for the lack of getting a mammography performed was that the physician did not order one.

One year after the 1977 release of the American Heart Association guidelines on antibiotic prophylaxis, a study found only 37% of patients eligible had received prophylaxis. Sagert et al. found only 22% of the patients got prophylaxis as
recommended by the practice guideline. The most common reason for the omission of prophylaxis was the lack of recognition of need by the physician and the selection of the wrong choice of antimicrobial agent or dose. Brooks et al. reported only 15% of dentists followed the same guidelines. Five years after the release of the American Academy of Pediatrics practice guidelines on the treatment of acute diarrhea, the authors found only 2% of physicians in the study followed the recommendation to rapidly rehydrate the children within 4 to 6 hours.

In a study that examined the National Cancer Institutes Community Hospital Oncology Program’s practice guidelines for staging and treatment of cancer, the authors report dismal results at best. Staging was recorded in 67% of the charts for lung cancer patients and only 33% of charts for breast cancer patients, respectively. Radiation therapy that was a recommendation of the guideline for the treatment of lung and breast cancer, was documented as ordered for 50% of lung cancer patients and 27% of breast cancer patients respectively.

The observational study results emphasize the need to implement practice guidelines better and more efficiently. Few behaviors are changed by simple dissemination of practice guidelines when used alone.

Behavior – Pre and Post Implementation Studies
A second group of studies on physician behavior include those in which a measurement is taken before implementation of the guideline and after implementation. This is the most popular form of evaluation of practice guidelines on physician behavior.

The State of Florida passed legislation mandating a practice guideline on Cesarean-section (C-section) be disseminated to all obstetric physicians. Each hospital was required to provide the exact date of implementation of the guideline. In a retrospective pre- and post-implementation study, the guideline did not accelerate the already downward trend in C-section rates. The conclusion of the study was that dissemination of a practice guideline does not achieve the magnitude or the specificity of the results desired. In a similar study, Gleicher et al. found that cesarean section rates continued to increase five years after the release of the National Institutes of Health Consensus Conference practice guidelines to reduce the use of these procedures. In the study by Lomas et al. presented earlier in this Chapter, the authors noted that the self-reported rates of change in the use of cesarean section did not match the rates found in actual practice.

Kelley et al. found a much more positive effect on behavior following the implementation of the American College of Cardiology practice guideline on the use of pacemakers. The study showed that following distribution of the practice guideline, there was a significant reduction in the use of the implants.

In a study of a guideline focused on patient length of stay (LOS) in chronic obstructive pulmonary disease (COPD), no reduction in LOS was noted. The reason
for the lack of effect was due to the fact that the guideline was overly restrictive with its inclusion and exclusion criteria. Only 19 percent of all of the patients diagnosed with COPD were included in the study. It should not be expected that 19 percent of a patient population would represent an entire homogenous population (i.e., all COPD patients). Even for the 19%, no change in LOS was noted for the treatment group as compared to controls. Therefore, no overall reduction in COPD patient LOS was noted. The authors conclude that although a practice guideline may appear efficacious, it may lack effectiveness when applied in clinical settings and may even increase costs.54

In another study, Weingarten et al. found adherence to a low risk chest pain guideline in the emergency department led to a reduced number of hospitalizations.6 However, adherence to the guideline also had an unexpected effect. If the guideline had been strictly followed, the institution would have been short of intensive care unit (ICU) beds for 26 percent of the patients treated under the guideline. Rigid adherence to the guideline would have increased or costs (increased utilization of ICU beds) yet had a reduced number of admissions (guideline goal). Thus, it is critical that guidelines be tested in a real clinical setting.

In a study of the AHCPR guideline which recommends the use of lumbar radiograph for the early diagnosis of low back pain, there was an increase in sensitivity to detect the occasional patient with low back pathology.119 Only 127 of the 963 (13%) patients with acute low back pain had a lumbar radiograph. Had the practice guideline been followed, 44% of the patients would have had the expensive diagnostic test (an
increase in utilization of 238%). The problem is that only 8 of the 963 patients had a
diagnosis of fracture or bone tumor. Thus, the cost of this utilization was not justified
due to specificity problems.

Faryna et al. monitored the outpatient prescribing patterns of internal medicine
residents and evaluated the effect of placing a one page set of antibiotic guidelines in
each patient examination room.\textsuperscript{120} The study found no differences in the rate of
inappropriately prescribed antibiotics due to the intervention. The authors concluded
that rapid availability of information about appropriate antibiotic use is not effective in
changing antibiotic choices.

Gupta et al. evaluated a locally modified practice guideline which address
intravenous analgesics, sedation and sustained neuromuscular blockade in adult
critically ill patients. National guidelines were adopted to local environment by a
consensus approach. The guideline was implemented by holding in-services and having
a pharmacist call the nurse and house staff that initiated an order for medication
involving sedation or agitation. The results showed a reduction in the cost of sedation
by 40% and neuromuscular blockade by 30%. The estimated annual cost savings were
$40,000.

The results of the pre and post implementation studies again show mixed results.
There has been little improvement in physician behavior as a result of practice
guidelines. These results do not necessarily mean the guidelines did not have any effect
on any clinician. The result does question the implementation methods in current use.
Behavior – Interventional Studies

The final group of behavioral studies involves those where a significant effort at implementation with various methods occurred. In a review of 102 educational intervention studies where different methods of implementation were compared, the authors found there was no method that proved universally to change behavior. The review concluded that dissemination only strategies were the least effective. However, even complex strategies ranged from ineffective to highly effective.

Schectman et al. disseminated a practice guideline on thyroid hormone testing. After the dissemination only method, compliance with the recommendations increased from 35 to 67%. Hence, dissemination does have some effect. Wachtel et al. had local physicians develop practice guidelines that gave recommendations on laboratory test ordering. Seventy-nine physicians were involved in the guideline development of the study, and this group generated 21 to 34% fewer laboratory test charges than in the previous year. However, even physicians not involved in the development had a 14 to 30% reduction in laboratory test charges. Marton et al. found that dissemination of a practice guideline focusing on the cost effective use of lab tests to medical residents had no effect.

Eisenberg et al. argues that behavior change by physicians require interventions which extend far beyond simple education. These authors favor interventions such as peer review, feedback, participation, penalties or rewards. Several studies have shown
that coupling practice guidelines with computer reminders, feedback and audits can significantly improve behavior change and impact outcomes.\textsuperscript{127-135}

Again, the results of implementation studies show extremely mixed results.

**Evaluation (Outcomes)**

The current section focused on literature that examined the following four outcomes: clinical outcome, economic outcome, health status and cost effectiveness. It has been stated earlier in this report that it is not sufficient to develop practice guidelines, they must be tested and effect reported. Practice guidelines should be outcomes based and justified. The worst outcome may be the result of adoption of a practice guideline that lowers the quality of care at a higher expense.\textsuperscript{38} Until outcomes are examined, this information may not be available. Most studies conducted to date concerning the effects of practice guidelines have looked at the process of clinical care rather than outcomes for patients.\textsuperscript{38} This assumes that improvements in the processes of care will eventually result in improved outcomes for patients.

**Clinical Outcome**

There is some direct evidence in the literature that practice guidelines improve clinical outcomes. Weingarten et al. conducted a study to examine the effect of length of stay practice guidelines for patients with total hip replacement, hip fracture or knee replacement on patient outcomes.\textsuperscript{136} The variables examined included health status (discussed further in the health status section), hospital readmission rate, return to the
emergency department, return to work, recreation activity and patient satisfaction. The intervention consisted of an intensive education program. The results found no significant changes inpatient outcomes with the exception that patient satisfaction declined in the intervention group for knee replacement. The authors concluded that the practice guideline did not negatively impact patient outcomes. For a length of stay guideline, this is a positive result.

For intravenous to oral switch, there is little evidence that defines medically necessary duration of parenteral antimicrobial therapy or the length of stay for patients hospitalized with pneumonia.6 The decision is often made arbitrarily by physicians and with uncertainty of when to switch.137 Weingarten et al. investigated the safety and effectiveness of a practice guideline that recommends early conversion of low-risk patients diagnosed with pneumonia.6 If the guideline would have been followed, the length of stay would have declined without a loss of quality as judged by medical chart reviewers. Other studies have documented that carefully selected patients treated with oral antimicrobial therapy have comparable outcomes to patients treated with parenteral antimicrobial therapy.138-141

Marras et al. studied guidelines for the empiric treatment of community acquired pneumonia (CAP) to assist in prescribing appropriate antimicrobials.142 Non-compliance with the guideline recommendations occurred most in suspected aspiration (a well documented complication of pneumonia). Overall, compliance was high with
the guideline without any implementation. The result showed that compliance with the
guideline had no impact on in-hospital mortality or length of stay. The authors conclude
that practice guidelines effect on outcomes remain unproven.

Barboni et al. found that relapse of bronchospasm in patients treated in the
Emergency Department and then discharged was reduced from 36 to 13% after the
distribution of a clinical practice guideline.\textsuperscript{143}

In a literature review to assess the evidence for the effectiveness of practice
guidelines in improving patient outcomes in primary care, it was found only 5 of 13
(38\%) of the included studies produced statistically significant results.\textsuperscript{38} In the 5 studies
with improvement, effect was noted for a proportion of the conditions studied, for only
certain subgroups of patients, or for a limited period of time. The conclusion of the
review is that there is very little evidence that the use of practice guidelines improve
patient outcomes in primary medical care.

In a review conducted in Britain, 59 published evaluations of practice guidelines
revealed that most studies showed the practice guideline had some effect on the process
of care (in all but 4 studies).\textsuperscript{58} However, only 11 (18.6\%) examined the impact of the
guideline on patient outcomes. Nine of 11 reported significant improvement in patient
outcomes. The conclusion of the review noted that although significant patient outcome
improvement was seen, the size of the effect varied.\textsuperscript{58}

The results presented above demonstrate the lack of literature support that
practice guidelines effect clinical outcomes. Yet, the main reason developers state for
the resources spent is to improve quality.
Economic Outcomes

There are several studies that have examined the effect of practice guidelines on economic outcomes. Marton et al. found that average laboratory charges for patients of medical residents who were given practice guidelines on laboratory use were $21.30 per patient visit. The laboratory charges for residents not given the guidelines were $31.10 per patient visit. Burns et al. found that locally developed practice guidelines on peptic ulcer disease decreased the costs for a health plan. The results were obtained by the prescription of less expensive medication therapies, fewer ordered endoscopy procedures and less frequent hospitalizations.

Gleason et al. retrospectively classified antimicrobial therapy for patients as consistent or not consistent with The American Thoracic Society's Guidelines for the Treatment of Community Acquired Pneumonia. The goal of the practice guideline was the appropriate prescribing of antimicrobials for empiric treatment of community acquired pneumonia. The guideline consistent group had reduced therapy costs for young patients (< 60 years old) with no co morbidity when compared to the not consistent group. However, the consistent group had higher therapy costs for young patients with co morbidities and the elderly when compared to the not consistent group. No difference between the groups was noted for any the clinical outcomes studied. The guideline sanctioned a clinical practice that was more expensive and gained no benefit for certain patient populations. Thus, no apparent benefit of the more costly practice guideline recommended therapy occurred in patients with comorbidity or old age.
In a modeling study, Cromwell et al. determined that AHCPR’s Smoking Cessation guideline was extremely cost effective. The guideline would cost $6.3 billion to implement and would produce a cost of $3,779 per quitter, $2,587 per life year saved and $1,915 per quality adjusted life years saved. The model assumes that physicians can identify all smokers and that all physicians are willing to counsel all smokers to quit. The study excluded lifetime medical costs. It is well established that the insurance industry has supported and implemented smoking cessation guidelines. However, few of the insurance companies that have smoking cessation guidelines provide coverage for the cost of smoking cessation aids such as nicotine gum or patches.

In a preventive care guideline, Stange et al. found that provider financial incentives did not favor using the guideline. Patients required longer office visits to provide preventive care services (10.9 versus 8.8 minutes on average).

In a prospective, interventional trial to study the effect of a length of stay practice guideline on patient outcome in patients with low risk with pneumonia the study found no significant difference in compliance or length of stay in the intervention group. The intervention consisted of providing the physician with patient specific risk information based on the practice guideline. There also was no impact on patient outcome. The author noted that patients already had a shorter length of stay than recommended by the guideline. Thus, there was no room for improvement.

One cautionary note in the literature is that previous studies have shown that cost containment efforts often ‘squeeze the balloon’ and simply shift costs from the inpatient
setting to the outpatient setting. The net result is not a net reduction in overall
cost. This is a significant concern for institution guidelines.

The result of the economic studies are mixed. Some show a decrease in cost
after a guideline is implemented and some did not find this effect.

Health Status and Cost Effectiveness

There has been few studies of practice guidelines on the health status as the main
outcome. The study presented earlier by Weingarten et al. used the SF-36 to measure
the effect of a length of stay guideline. No effect on health status was found. Since
the purpose was to show that length of stay could be reduced without a negative impact
on health status, the result is indeed a positive one. In a similar study, Gleason et al.
also used the SF-36 as an outcome measure to show that compliance with prescribing
antimicrobial has no negative impact on health status. Again, no effect was correctly
interpreted as a positive result.

In a cautionary note, Granata et al. noted that most clinicians value the
optimization of individual health outcomes. Delivery of the most cost-effective care for
groups of patients via a practice guideline may require different management
strategies. Practice guidelines to promote cost-effectiveness at the patient level may
not maximize cost effectiveness at the population level.

In a survey of insurance representatives, employee health benefit managers and
corporate medical personnel report that one of two main needs was that of determining
the impact of practice guidelines on health status. There has not been significant study of the cost effectiveness of complying with the recommendations of a practice guideline.

**Evaluation Barriers**

At an institution level, it has been argued that a maximum of 12 to 15 guideline efforts should be undertaken at any one time. By focusing on the vital few, evaluation becomes possible. The tendency within a single institution is to have as many guidelines as possible. The problem quickly becomes one of quantity versus quality.

The single largest barrier to the evaluation of practice guidelines is a lack of resources. First, so much time and financial resources are devoted to development, little is left for evaluation. Second, the complexity of the methodology requires a great deal of effort not normally employed in medical research. Third, there is a lack of data to evaluate. The result is produces a need to do primary data collection which is time consuming and expensive.

The development and implementation barriers presented earlier also tend to limit evaluation. As noted, often guidelines do not present the objective or goal it is intended to impact which makes it difficult to evaluate. Often the wording or language of practice guidelines is highly specific to achieve clarity. However, the scientific validity of such wording is often questioned by practicing clinicians.
Evaluation Summary

It is dangerous to view something as a standard of care unless outcomes and patient preferences are known. To do so implies health science knows more than justified. This underscores the need to study the outcomes associated with practice guidelines.

Many in the health care industry recognize the importance of evaluation of practice guidelines, yet most continue to put the majority of resources into development. Research is clear – the mere existence of practice guidelines does not guarantee they will change behavior or influence outcomes. More research attention needs to be applied to evaluation of practice guidelines.

It has become imperative that the effects of clinical practice guidelines on outcomes be evaluated. Evaluation must be done to determine which practice guidelines are beneficial and to prevent the codification of low quality practice guidelines.

Practice Guideline Literature Summary

Health care practitioners with similar clinical experience and education will select different treatments and tests for patients with similar medical conditions. Significant variability in health care exists in the United States that cannot be accounted for on the basis of severity of illness, patient demographics or clinical necessity. Physicians tend to formulate individualized patient care strategies in belief that each clinical situation is unique. Decisions are guided by the training, clinical experience and
knowledge of the provider in conjunction with patient preferences. An approach to promote appropriate decision making is to make the most recent and relevant information available to providers via a practice guideline. The issue is whether or not this medium, despite large resource investments, is used in practice for the benefit of patients. The belief that clinical practice guidelines can improve the quality, value, and effectiveness of health care has resulted in proliferation of practice guideline efforts. However, much more research needs to be conducted to determine if the hope is fulfilled.

There are several general characterizations of the literature that are important for the purpose of this study. First, practice guidelines have become tools used to state the standard of care for many treatments. Second, there are several purposes and uses for guidelines. Yet, few studies examine the cost and effectiveness of guidelines. Third, most guidelines are initially developed at a national level and modified to 'fit' local clinical practice or developed at the local level. The impact of local guidelines has not been sufficiently studied in the literature. Fourth, compliance with guideline recommendations is variable. Uncertainty in practice includes following or not following guideline recommendations. More importantly, there are few studies in the literature which link guideline compliance to patient outcomes. Finally, there are few studies of guidelines that are conducted locally or study the cost and effectiveness of compliance with recommendations.
General Concerns with Practice Guidelines

Upon reviewing the literature on practice guidelines, a final section is necessary to understand the general concerns within the medical community. A major concern is that practice guidelines may promote ‘cookbook medicine’. Medical educators fear that early exposure to practice guidelines may even discourage young physicians from acquiring skills in clinical reasoning and limit independent thinking through the logic of therapeutic choices.

A second concern is that in the effort to reduce variation, unhealthy uniformity in medicine may occur. Uniformity may not respect true differences in patient populations or practice settings. Researchers fear that practice guidelines may discourage innovation. Funding agencies may not be willing to award grants for studies of treatments that are not supported in a practice guideline because of a lack of evidence.

There is significant concern about how to keep practice guidelines current in an environment of rapid change. Some physician subspecialty groups may feel and actually be threatened by practice guidelines developed by another group. Practice guidelines may indeed fuel ‘turf wars’ to defend rights to perform and bill for procedures. A concern noted throughout this Chapter is a threat to autonomy. Eddy et al. noted “whoever controls practice policy controls medicine.” There is a real fear of unreasonable recommendations prepared by nonclinicians unfamiliar with patient care.
The medical community has expressed concern over the potential uses of practice guidelines. Guidelines may be used by payers as a basis for denying coverage. Guidelines may be tied inappropriately to credentialing and hospital privileges for physicians. Guidelines may be misused as evidence in malpractice cases.

Patients may be disappointed when a practice guideline blocks access to desired services, providers or coverage. Guidelines designed to optimize outcomes for society may not meet the needs of individuals. Practice guidelines may be a source of health care rationing. Finally, payers and politicians worry that practice guidelines may not reduce cost. By promoting services that are underutilized, a guideline may increase cost.

It is time more studies are done to address these concerns and provide a basic reasoning and justification for the massive production of clinical practice guidelines.

HEALTH STATUS

In the previous discussion of health status, only studies which used health status and focused on practice guidelines were presented. The focus of this section is to provide a brief review of the general health status literature. In the past thirty years interest in the measurement of subjective health status has blossomed. The result of heightened interest is the creation of a massive number of scales and instruments available for quantifying subjective health status. The purpose of this section will not be to provide a complete review of this literature. Rather, the purpose is to discuss
some of the basic issues and topics relevant to this study and the use of a health status measurement tools. The first part of the review will focus on broad topics within the literature such as the uses of health status measures and selection criteria for use in a study. Second, the section presents some of the psychometric properties needed to use a health status measure. Finally, the review will focus on the SF (Short Form) 12 as the instrument for use in this study.

Uses of Health Status Instruments

Health related quality of life (HRQOL) refers to the experience and importance of different domains of health that are affected by disease and treatment.\textsuperscript{174,175} Five concepts may be used to define HRQOL including 1) impairment, 2) functional states, 3) health perceptions, 4) social opportunities, and 5) duration of life.\textsuperscript{176} Domains or health constructs are operationalized to translate these five concepts of quality of life into measurable terms.\textsuperscript{177} There has not been a consensus formed on the optimal inclusive set of domains which need to be operationalized. This is no surprise due to the fact that quality of life has different meaning for different people. Kane et al. lists the following as the most common domains used in instruments: physical functioning, social functioning, emotional functioning, cognitive functioning, pain, vitality and overall well-being.\textsuperscript{177} Some instruments will have all or some of these domains. The key issue is selecting an instrument that meets the needs of the research population.

There are three types of health status instruments presented in the literature based on the intended population for whom the instrument was designed. The first type
of instrument is the 'generic' instrument. A generic health status instrument is designed to be used in a heterogeneous population such as a general population, a sample of patients with different diseases or a sample of patients with different treatments.

Generic instruments have an advantage in generalizability of the results. Generalizability allows generic instruments to be used across different disorders, severities of disease, interventions, and demographic and cultural groups. The generic instrument is also well suited to measure health status of patients with physical and mental conditions and compare the results to other populations such as a 'normal' healthy group of patients. Examples of generic health status instruments include the Functional Status Questionnaire, the Dartmouth COOP Poster Charts, the Nottingham Health Profile, the Duke Health Profile, the MOS Short Form 36 (SF-36), the MOS Short Form 12 (SF-12) and the Sickness Impact Profile. This is not a complete list of generic instruments but a sample of available instruments.

The second type of health status instruments are termed 'disease specific' instruments. These instruments are targeted at distinct populations of patients such as patients with asthma, arthritis, depression, epilepsy, etc. Disease specific instruments are designed to measure items that are of importance to the targeted population. Unlike the generic instruments, the disease specific instruments avoid the collection of information that may be irrelevant. Disease specific instruments are usually more sensitive measurement and are designed to capture small changes in patient
functioning. Sensitivity will be discussed further in the psychometric section below.

Examples of disease specific instruments include the Asthma Quality of Life Questionnaire and The Beck Depression Inventory.

The final type of health status instrument are termed 'patient-specific scales'. In these types of instruments, no two people have exactly the same scale. Each patient selects the most troubling symptoms. The scale is constructed and scored over time. Thus, the patient serves as his or her own norm for scoring. These instruments have the distinct advantage of measuring not only domains important to the diseased population but also of each individual patient. An example of a patient specific scale for heart failure patients is provided by Guyatt et al.

For the purpose of this study, a generic instrument was selected as the most appropriate instrument type. There is not a well defined clinical patient population in this study. The patients all received antibiotic therapy. However, many have different types of underlying disease processes. For example, there are patients in the study who have had myocardial infarction, lung embolism, transplant complications, diabetes and a variety of primary infectious disease processes. In other words, the patient population for the study is very heterogeneous. Therefore the remainder of the section will focus on generic types of health status measurement instruments.

Instrument Selection Criteria

Kane et al. provides a general list of criteria for the ideal health status measure. The criteria include: domains of health, range of health, clinical relevance, level of
emphasis, sensitivity, reliability, validity and practical considerations. The selection of
the instrument for use in this study was based on the following criteria and will be
discussed further: 1.) domains of health, 2) range of health, 3.) clinical relevance, and
4) practical considerations.

The domains contained within an instrument were an important consideration
when selecting a generic health status instrument. Due to the heterogeneous population
of patients in this study a wide range of dimensions was needed to measure and quantify
health status. Physical functioning, mental health, social functioning and pain were
targeted for measurement in the study. Many of the published generic instruments
contain scales that measure these domains. Thus, this was not a criteria that was
in and of itself sufficient to select an instrument.

Next the range of health to measured by the instrument must be large. Range in
health refers to the possible stages of dysfunction that patients in a study may
experience. Patient health lies in a continuum that spans from perfect health to death or
even worse than death. For the purpose of the study a wide range of health was
needed in the measurement since some patients would be very ill and others not so ill.

One issue that was considered was the ‘ceiling’ or ‘floor effect’ involved in
measurement. All measures have floor and ceiling effects. A floor effect occurs
when an instrument is not able to measure health below a certain point on the scale. A
ceiling effect is the inability to capture health above a certain range. Since many of the
health status measures have these effects, this criteria was not a major one in the
decision process. However, as will be explained in Chapter 3, these effects were
examined.

The third criteria, clinical relevance, is an important issue to this study. Clinical
relevance refers to the usefulness of the information derived from the measure.\(^{177}\) A
related issue is whether a researcher is interested in measuring at a population or group
level or at an individual patient level.\(^{190}\) Health status measures have been utilized in
many group-level applications such as: describing health profiles for patients differing
in diagnosis, disease severity and treatment regimen, evaluating the relative benefits of
different treatments, comparing health outcome across different health care delivery
systems, assessment of health policy initiations and measuring general population
health.\(^{190}\) Interest has also grown in using health status measurement in clinical practice
for individual patient assessment and monitoring.\(^{190}\) In the later use, the unit of analysis
is an individual patient rather than a group of patients. McHorney et al. cautions that of
five generic instruments examined, none were precise enough for individual patient
monitoring.\(^{190}\) For the purpose of this study group level profiles were the reason for
measuring health status. No individual health status measurement is utilized. Thus the
clinical relevance of the information obtained is low.

The final criteria and a major one is practical considerations. Practical
considerations include issues such as the length of time required to administer the
questionnaire, the appropriate format for the survey, the cost of administration, the
complexity of the measurement and scoring methods, and the acceptability of the survey
to patients. In this study, patient burden considerations were a significant consideration. Due to the severity of illness in the patient population, the number of items included needed to be kept to a minimum. The format most appropriate for administration was a survey based instrument. Total time to administer the survey also had to be kept to a minimum. Few instruments in the literature meet these practical criteria. The 12 item SF-12, the 9 item Dartmouth COOP Poster Charts and the 17 item Duke Health Profile were candidates. Based on practical considerations, the SF-12 was selected as the health status instrument for the study. The rest of the discussion will focus on the SF-12 generic health status instrument.

**Psychometric Considerations**

This review will not cover in depth the vast literature on psychometric theory. Any reader interested in reading more about this topic is referred to Nunnally et al. The review will focus on published studies of the psychometric properties associated with the SF-12. The development of the instrument will be presented along with the psychometric property descriptions and definitions. Properties of the instruments discussed will include the reliability, validity and responsiveness. Each discussion will review the impact of the specific information on the purpose of this study.

The basis for the development of the SF-12 was practical considerations. There are many studies in which the SF-36 and other generic instruments are too long for practical use. The SF-12 was developed to provide a shorter instrument for use in large studies that would take less than 3 minutes to complete. The issue is whether or
not the psychometric properties of the SF-12 are good enough to make the instrument useful in practice. Two main psychometric factors of the SF-36 led to the development of the SF-12. First, physical and mental health factors account for 80 to 85% of variance in all eight of the SF-36 scales. Second, the SF-36 Physical Component Summary (PCS-36) captures the hypothesized differences on physical criteria and the Mental Component Summary (MCS-36) captured hypothesized differences in tests using mental criteria. Thus, it was possible to reduce the number of items without a loss of much information. In this study only the two summary scores on physical (PCS-12) and mental (MCS-12) were needed which makes the SF-12 a suitable instrument for consideration. The SF-12 summary measures were constructed to reproduce the SF-36 physical and mental summary measures. The SF-12 also has items in it, which represent the eight SF-36 domains. The domains (with the number of items in parentheses) include physical functioning (2), role-physical (2), bodily pain (1), general health (1), vitality (1), social functioning (1), role-emotional (2), and mental health (2).

Scoring of the SF-12

The SF-12 utilizes norm-based, standardized scores, which compute a Physical Component Summary (PCS-12) and Mental Component Summary (MCS-12), in that order. Physical, and mental regression weights and a constant for each summary were derived from the U.S. general population. The transformed or standardized score results in a mean of 50 and a standard deviation of 10 in the general U.S. population. It is
important to note that the SF-12 scores on the two summary scores are computed relative to the general U.S. population. This was done by the instrument developers to aid in meaning. The scores may be meaningfully compared with the other and have a direct interpretation in relation to the distribution of scores in the general U.S. population. The mechanics of computing the actual scores is provides in the manual for the instrument and will not be reviewed further.  

Reliability

Measurements are reliable to the extent that they are repeatable and that any random influence, which causes different measurements of the same variable to vary, is a source of measurement error. There are two distinct operational definitions to test reliability. The first are tests of internal consistency or tests of homogeneity. Examples of this type of test include Cronbach’s α and Kuder-Richardson formula 20. The basis for these test is that the expected correlation of one test with an alternate form containing the same number of items. These tests of reliability are estimates based on the average correlation among items within the scale or instrument. The second definition may be termed temporal stability. In this form of reliability testing the same individual may be given the same instrument separated by time (test-retest). Correlation of the items are then computed. It is important to note these two forms of reliability testing are not the same. One could have high test-retest reliability (correlation) and yet have low internal consistency due to low item correlation. Reliability is improved by
making items and instructions clear, making test administration consistent, increasing the number of items on a test, reducing ceiling and floor effects and increasing item correlation.

For the SF-12, test-retest in a general U.S. population reliability coefficients have been reported in the literature. Ware et al. reports that the test-retest reliability for the PCS-12 scale is 0.89 and for the MCS-12 is 0.76.\textsuperscript{194} Hurst et al. reported test-retest reliability in patients with rheumatoid arthritis for the PCS-12 was 0.75 and for the MCS-12 was 0.71.\textsuperscript{195} Reliability coefficients of 0.70 or greater are generally satisfactory for scales used in group-level analysis.\textsuperscript{192} The internal consistency of the SF-12 was not reported. Based on the test-retest results, the SF-12 has minimal support for reliability. It should be no surprise that there is some loss of reliability as compared to the SF-36 due to a reduction in the number of items. This study will examine the reliability of the SF-12 in the population of patients. Health status instruments should be examined for reliability when utilized in a new patient population. This study will examine the internal consistency of the SF-12.

**Validity**

Validity refers to the extent to which a measure actually measures what it says it does and is not measuring something else.\textsuperscript{173} One significant issue is that reliability and validity are not independent concepts. Reliability is a necessary but not sufficient
condition for validity. No instrument is proven valid for use in any specific population. Empirical testing of validity is necessary and should be undertaken for all studies using health status instruments.

There have been traditionally three types of validity: content, criterion, and construct. Content validity assesses if the items of an instrument are representative of the content to be measured. For example, the PCS-12 should appear to measure physical functioning. Criterion related validity is when the scale score is correlated with some other measure of the same construct under study. Ideally, a ‘gold standard’ that is widely accepted in practice would be used for criterion validity testing. In health status measurement no perfect ‘gold standard’ exists. However, many studies use a criterion that is close to the construct and correlate the instrument score with it. The third type of validity is termed construct validity. There are four kinds of testing done for construct validity: 1) extreme group, 2) convergent group, 3) discriminate group, and 4) multitrait-multimethod matrix. An extreme group consists of a population ‘known’ to have a trait or characteristic. A validity study compares the extreme group score to the group without the trait or characteristic. The instrument should show a difference in scores for the two groups. Convergent refers to the fact that a measure should correlate with other measures of the same construct. Discriminate refers to the fact that a measure should not correlate to measures of different constructs it is not supposed to measure. The fourth kind of test for establishing evidence for construct validity will not be discussed in detail. The basic description for multitrait-multimethod matrix is to give the same instrument and other similar instruments to the same population using
two different methods of administration. The result is a series of correlation matrices that are interpreted for support of construct validity.

For the SF-12, content validity appears to be acceptable. All of the items on the SF-12 are taken directly from the SF-36 that has been tested for reliability and validity in several diverse populations.

Ware et al. uses extreme groups to test the validity of the SF-12. The PCS-12 and the MCS-12 have been shown to discriminate between four defined groups. The four groups were known to be different in physical and/or mental health as defined by clinical measures. The criteria used to define groups is not relevant to this study and will not be discussed further. The results of the validity testing for the SF-12 were also compared to identical studies done for the SF-36. The results showed that in tests of physical differences, the PCS-12 was a better predictor than the MCS-12 as expected. In tests of mental differences, the MCS-12 scores were better predictors than the PCS-12 scores as expected. The authors conclude that the SF-12 loses some validity by reducing the number of items as compared to the SF-36. However, the conclusions reached using the SF-12 were the same as those reached using the SF-36.

Responsiveness

Responsiveness or sensitivity to change refers to the instruments ability to detect clinically important changes over time, even if those changes are small. There has not been significant testing of the responsiveness of the SF-12. Hurst et al. conducted a study comparing the SF-12 to the SF-36 in patients with rheumatoid arthritis. The
results of the study used the standardized response mean (SRM) to show that the SF-12 was slightly less responsive than the SF-36 to improvements in physical health (SRM = 0.52 for PCS-12 and 0.61 for the PCS-36). The SF-12 was also less responsive in mental health improvement in this patient population as compared to the SF-36 (SRM for SF-12 = 0.31). This result for the MCS-12 correlates well with the results obtained from the SF-36.195

Health Status Summary

Although there is some loss of reliability, validity and responsiveness, the SF-12 appears to be suitable for comparing groups of patients. It appears to be applicable to groups of patients who have large changes in health. The psychometric properties of the SF-12 were tested in this study.

COST EFFECTIVENESS

Health care economics is a study of the allocation of scarce health care resources. There are many approaches for the evaluation of alternatives from an economic perspective. Hakim et al. notes the reasons for this popularity of economic evaluation are: increasing need for accountability, making decision process more explicit, evolving capitated payment systems, increasing high cost technology development and rising health care inflation.196 Economic evaluation techniques by definition assume that resources saved will not be wasted but be put to efficient use.197
Pharmacoeconomics may be defined as the comparison of outcomes and costs of a pharmaceutical program or service to alternatives from specified perspectives. The goal is to link inputs used in producing outputs (outcomes) of producing programs and services. Sackett et al. provides the four ‘E’s of economic evaluation. The first ‘E’ is efficacy or can a treatment work under ideal conditions. The second ‘E’ is effectiveness or does a treatment work under real conditions. The third ‘E’ is equity or is the treatment reaching these individuals who need it. The final ‘E’ is efficiency or is the treatment worth doing. The focus of this review is on the second ‘E’ or effectiveness.

Hakim et al. outlines the steps involved when conducting a pharmacoeconomic (PE) analysis. The review will discuss each step in context with the current study. The first step in an economic analysis involves the definition or identification of two or more alternatives for comparison. For the present study, the alternatives are writing a practice guideline compliant order or a non-compliant order.

The second element of a PE evaluation is the costing of the alternatives. Costing includes the identification, measurement and valuation of all costs involved in the study. A significant issue is the perspective of the study and this must be explicitly stated. Many costs are revenues for another perspective in most industrial systems. In this study, the perspective of the institution will be taken. Since the practice guidelines were developed and implemented in a single institution, the institutional perspective was felt to be the most appropriate.

The third step involves the measurement of the outcomes of the alternatives. Kozma et al. described the Economic, Clinical and Humanistic Outcomes (ECHO)
model. Economic outcomes may be direct or indirect benefit. It is important to make a distinction between cost and benefit. In the present context the term economic refers to benefits. Direct economic benefit are the savings in treatment costs and indirect economic benefits are measured by calculation the reduction in production losses due to a treatment alternative. Clinical outcomes include natural biologic units such as mortality or morbidity. Clinical outcomes may be sub classified as final or intermediate outcomes. For example, when comparing blood pressure treatment alternatives a final outcome may be reduced rates of death or heart attacks. An intermediate outcome may be the lowering of blood pressure in mmHg. Humanistic outcomes include health-related quality of life and patient satisfaction.

The fourth and final step in a PE evaluation is to link the costs and outcomes for the alternatives. There are four full economic evaluation techniques as summarized by Table 2.2. The most distinguishing features of full economic evaluation are that they all compare two or more alternatives and both costs and consequences are examined. The difference between the four types of economic evaluations are the units of which the measurement of outcomes or consequences are measured and reported. For the purpose of this study, CEA was selected as the method of choice. CEA is useful since the cost of care for patients were readily available. The Ohio State University Medical Center has an accounting system in place that estimates direct cost data and allocates indirect and overhead costs to each patient care episode via the use of cost to change ratios. However, utility and benefit data is not readily available. Effectiveness has been
defined as successful antimicrobial therapy (previously defined). Thus, the enumeration, measurement and valuation of the costs and outcomes were suitable for CEA.

Cost effectiveness as a term has been quite popular in the literature.\textsuperscript{202} Unfortunately, many studies that use the term fail to provide the required data.\textsuperscript{202} The key is the linking of the costs and the outcomes. Knowledge of the outcomes and the costs of alternatives is insufficient to label an alternative as cost effective.\textsuperscript{202} Doubilet et al lists four common errors for labeling an alternative as cost effective made by authors in the literature.\textsuperscript{202} First, that cost effective equates to cost saving. Under this criteria alone, many current treatments and therapies would not be done since many do not save money but provide a benefit. To save money, medications could not be used at all. However, it would be ridiculous to claim this as cost-effective. Second, cost effective equates to effective. This scenario occurs when cost is not considered. This is an efficacy versus effectiveness issue and is misleading to readers.\textsuperscript{199} Third, cost effectiveness equates to cost saving with an equal (or better) health outcome. This would eliminate the decision to do an alternative treatment that is much cheaper but has a slightly reduced effect. Finally, cost-effective equates to having an additional benefit worth the additional cost. This is a value judgment that is wholly dependent upon one’s perspective.
<table>
<thead>
<tr>
<th>Type of Study</th>
<th>Costs</th>
<th>Consequences</th>
<th>Measurement of Consequences</th>
<th>Compare Alternatives</th>
<th>Assumes Equivalent Effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Benefit Analysis (CBA)</td>
<td>Monetary value of resources consumed</td>
<td>Monetary value of outcomes</td>
<td>Economic</td>
<td>Not necessarily although comparisons are implicit</td>
<td>No</td>
</tr>
<tr>
<td>Cost Utility Analysis (CUA)</td>
<td>Monetary value of resources consumed</td>
<td>Utility of health effects</td>
<td>Quality Adjusted Life Years (QUALYs)</td>
<td>Economic</td>
<td>Economic</td>
</tr>
<tr>
<td>Cost Effectiveness Analysis (CEA)</td>
<td>Monetary value of resources consumed</td>
<td>Indirect Costs Subsequent use of resources</td>
<td>Effects on health</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Cost Minimization Analysis (CMA)</td>
<td>Monetary value of resources consumed</td>
<td>Direct medical costs</td>
<td>Clinical outcome</td>
<td>Economic</td>
<td>Yes Yes</td>
</tr>
</tbody>
</table>

* Adopted from unpublished notes of Pathak DS.  

Table 2.2 Full Economic Evaluation Techniques
Some criterion for cost effectiveness is crucial for meaning. Doubilet et al.
recommends that the term cost effective be restricted to the fourth error above when a
full explanation of how the 'worth' is spelled out. Three scenarios lead to this
conclusion. First, an alternative is cost effective if it is less costly and is at least as
effective or more effective than an alternative. Second, an alternative is more effective
and more costly, but the additional benefit is 'worth' the added cost. Third, an
alternative is cost effective if it is less effective and less costly with the added benefit of
the alternatives not being 'worth' what it costs. The second and third scenarios require
a utility or value be placed on the alternative. The current study did not do that. Only
the first scenario will be considered cost effective in this research.

There are three common decision rules used in CEA. The first is the
domination rule. Any alternative that is dominated should be eliminated as an
alternative. The second rule is the budget as the decision rule. After an incremental
analysis has been done, a fixed dollar budget is used to implement the most cost
effective alternative and proceeding to the next alternative until the budget is exhausted.
For the purpose of this study, the fixed budget decision rule did not apply. The third
rule involves using the price per effectiveness unit as the decision rule. This rule allows
the judgment of the maximum amount willing to be paid for a unit of effectiveness. For
the purpose of this study, no set price was available. To set a price a willingness-to-pay
approach might need to be utilized. Thus, for the purpose of this study if an alternative is dominant, it was termed cost effective. Otherwise the alternative did not make a claim to be cost effective.

The underlying rationale behind CEA is the optimization or maximization of the health outcomes for a given amount of resources. Cost effectiveness was defined in the study as the present value of the cost of treatment per hospital stay over the rate of antimicrobial success (e.g., to produce a cure or prevention of infection). This calculation provides average cost effective ratios indicate less resources are used to produce a given outcome. Thus lower ratios were preferred. Clinical data were collected via a chart review. Cost data were accessed via the hospital's financial accounting system.

Two issues in CEA that should be discussed are timing effects and sensitivity analysis. When economic cost or benefit data is collected over a long period of time (> 1 year), then the monetary value should be discounted to reflect the present value. Due to the short duration of the time period for this study (less than 60 days), no discounting was employed. Sensitivity analysis is a technique to study the effect of uncertainties involved in the valuation of costs or consequences in CEA. In this study, the outcome (antimicrobial success) were altered 5.0% in either direction and CEA ratios recalculated. If none of the results changed, the results were considered robust to the assumption of the analysis.

Average cost effective ratios are of limited usefulness. The ratios do not compare the costs and outcomes of the alternatives directly. For example, an alternative
may have a low average cost effective ratios by having low cost and low effect or by having high cost and high effect. The magnitude of the differences between alternatives may not be readily seen with average CEA ratios. An incremental analysis compares the excess cost of one alternative over another to the excess effects it provides. The incremental cost effective ratio provides a direct comparison of the alternative to one another. An incremental analysis was performed in this study.

The results of the CEA may be presented in two forms — tabular and graphic. In the tables, the costs, outcomes and average cost effective ratios were presented. The data will not be provided in a decision tree format or a graphic format. Due to the comparison of only two alternatives with two outcomes, tables were deemed sufficient for data presentation. In separate tables the incremental CEA will be presented.

Any alternative which is higher in cost and lower in providing effect or outcome (infection cure in this study) is said to be dominated. A dominated alternative is eliminated and a new incremental analysis will be performed. A dominated alternative should never be done since it would never produce positive incremental or marginal gain in cost or outcome. Cost effectiveness planes will be graphed to show the data in graphical form.

Interpretation of results may show that following guidelines and not following guidelines are each equally effective. That is the purpose of this study.
REFERENCES FOR CHAPTER 2


119


120


194. Ware JE, Kosinski M, Keller SD. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales. Boston, MA; The Health Institute, New England Medical Center: 1995.


CHAPTER 3

METHODS

The goal of health care delivery is to maintain or improve health for patients. Measurement determines if processes of care provided to patients achieve positive outcomes (outcome measurement) or positively impact processes thought to be associated with good outcomes (process measurement).

To make judgments of value with respect to the impact of a delivery processes on outcomes, patient demographic and clinical history of patients had to be measured or accounted for. Variation and effects due to patient characteristics or clinical history were examined when measuring the impact of drug therapy compliance with guideline recommendations on patient outcomes or cost and effectiveness.

Process(es) of care may be characterized by what happens to the patient (i.e., diagnostic tests, procedures, treatments, etc.). For the purposes of this study, the only process measures utilized were for the determination of compliance with the guideline
recommendations. The rest of the variables used in the study were patient outcomes. Patient outcomes were measured and operationalized include: economic outcomes, clinical outcomes and health status.

Each of the two guidelines were evaluated independent of one another. The methods used for each guideline evaluation were similar. Thus, in the methodology section, only one discussion was presented. The two guidelines were selected based on the premise that they were developed to address identified clinical issues at OSUMC and each focuses on drug therapy.

This Chapter reviews the methods used in the study. First, the study design is presented. Second, the study presents a review of the non-response analysis utilized in the study. Third, development and evaluation methods for the data collection tool and measurement instruments are discussed. Fourth, the target population and sampling methods are reviewed and discussed. Fifth, the statistical analysis for the study is discussed. Finally, the Chapter ends with a discussion of selection bias and potential measurement error that may impact the study conclusions.

RESEARCH DESIGN

An observational, two-group, historical cohort study design was utilized for each of the practice guideline evaluations. This study design was the most appropriate since no experimental design was possible. It has been noted in the literature that randomized trial, may not be a 'gold standard in behavioral research.' There was no
controlled assignment (e.g., randomization) of patients to groups or randomization of a treatment. No control group was utilized in the study. Under the premises 'that guidelines are beneficial', there were ethical reasons to avoid using a control group. Since each of the two practice guidelines had already been implemented prior to the start of the study, no randomization was possible.

A cohort study is one in which cases were selected on the basis of a particular disease, exposure or treatment. The cases were then observed for effect. A cohort is defined as a group of individuals that share a common experience within a defined time period. The observational design allows the research to occur in a natural rather than controlled setting. The design helped answer the question of 'what happens in a real practice setting'. This type of design is useful for effectiveness studies. The two cohorts utilized in the study were 1.) patients treated in the guideline compliant order group and 2.) patients treated in the guideline non-compliant group.

The observational, cohort design was well suited for the research questions of the study. The natural decision of following a guideline recommendation or not following a guideline the recommendation was allowed to occur in a real clinical setting. The research questions posed in the study examine the impact of order compliance on patient outcomes. One significant disadvantage of the design was the need to follow-up all patients. Lost to follow-up of patients had the potential to bias the results of the study. This issue was discussed in detail in the target population and sampling section.
**Study Treatment**

The study did not implement or impact any treatment decision. The study measured the effect of treatment decisions made for particular sets of patients (i.e., patients who had drug therapy covered by one of the two guidelines). Observational data were used to define which cohort of patients had better outcomes.

**Data Collection**

Patients were enrolled in the study once an order for one of the guideline covered drug therapies was written. The patient was enrolled within twenty four hours after the order was written. The patient was selected randomly from a list generated by the Department of Pharmacy. Drugs covered by the respective guideline are listed in Table 3.1. Thus, once the medication order was received in the pharmacy, the patient was approached to participate in the study.

<table>
<thead>
<tr>
<th>Criteria for Use of Vancomycin Guideline</th>
<th>Converting Intravenous to Oral Antimicrobials</th>
</tr>
</thead>
<tbody>
<tr>
<td>vancomycin</td>
<td>ciprofloxacin</td>
</tr>
<tr>
<td></td>
<td>azithromycin</td>
</tr>
<tr>
<td></td>
<td>clindamycin</td>
</tr>
<tr>
<td></td>
<td>fluconazole</td>
</tr>
<tr>
<td></td>
<td>ampicillin/sulbactam</td>
</tr>
</tbody>
</table>

Table 3.1: Medication Therapies Covered By Guidelines
The study consisted of two measurement points in time, a baseline or initial measurement and a second follow-up measurement. The two measurements were administered approximately 30 days apart. Thirty days was selected since the main complication of the therapies (e.g., re-infection) would likely have occurred. Also, a shorter time period was selected to reduce the number of patients lost to follow-up. A cover letter explained the research, a patient consent form and the baseline survey was given to the patient in the hospital. Appendix D contained a cover letter utilized. Appendix E contained a patient consent form and Appendix F contained the baseline health status survey. The researcher explained the purpose of the research to the potential enrollee. If the patient agreed to participate in the research, the patient was asked to sign the patient consent form and to complete the baseline survey. If the patient did not choose to participate, the researcher recorded the reason given. Thus, enrollment into the study and initial data collection were done while the patient was in the hospital.

The survey was a written questionnaire that contained the baseline health status measure and basic descriptive questions concerning recent utilization of health services (see Appendix F). Written questionnaire format was used to keep respondent burden low and reduce the amount of resources needed to collect data. One significant advantage of the written format was lower cost and researcher burden. One significant limitation of the written format was the lack of ability of some respondents to complete the tasks of reading and writing. In this study, some patients asked to have
the survey read to them and then their answers were recorded on the instrument. The researcher used the instructions provided in the health status instrument manual.7

After the patient was enrolled in the study and had completed the baseline questionnaire, the patient's chart was reviewed to obtain patient documented data. This data was recorded on the patient chart review form that appears in Appendix G. The data collected from charts included process variable data and outcome data. The process data was used in the study to place a patient in a given cohort group.

Approximately, one month after the patient was enrolled in the study and had completed the baseline survey, the one-month follow-up survey was sent to the patient using the U.S. mail service. The one-month follow-up survey contained the same written health status questionnaire as the initial baseline measure. In addition, the follow-up survey contained resource utilization since the time the patient was discharged from the hospital (e.g. over the 30 days after discharge) and a question asking the patient to give a global rating of change in health status over the 30-day time period. The one-month follow-up survey instrument is presented in Appendix H.

Figure 2 contains a flow diagram of data collection process used in the study. All data from the chart review form and the two survey instruments were entered into a database for data analysis. Each of these instruments will be discussed in detail in the Instrument Design section.
Patient Confidentiality

All patient data were kept strictly confidential. The research was submitted and approved by the Biomedical Sciences Institutional Review Board at The Ohio State University (see Appendix I). Only aggregate data was reported in the study. After adequate explanation of the purpose of the study, each patient who agreed to participate was required to sign a patient informed consent form. The patient consent form is provided in Appendix E. All enrolled patients were assigned a 5-digit random identification number. This number was utilized to ensure patient confidentiality. For data collection purposes, patient names, addresses and 9-digit medical record numbers needed to be recorded until all data was input into the data base for each respondent.
After the patient data was entered into the database, all patient identifying information was kept strictly confidential. All signed patient consent forms will be kept for three years then destroyed.

GUIDELINE EVALUATION

Using similar items and methodology as reported by Shaneyfelt et al., the qualities of the two guidelines were assessed according to the standards presented in Table 1.1. Two practicing physicians were asked to rate the two guidelines independently from one another on these criteria. The physicians were asked to respond yes or no if the respective guideline met the development criteria. Each physician was provided a copy of the respective guideline and the Shaneyfelt et al. article for reference. The results were calculated as a percentage of standards met and compared to the results obtained by Shaneyfelt et al. The evaluation tool is presented in Appendix J.

NONRESPONSE AND RESPONSE RATE

This section of the Chapter focuses on two main issues. First, the nonresponse for the study was examined to determine if there is reason for concern when interpreting the results of the study for differences between patients eligible for enrollment into the study and those not eligible. Second, the section will examine the comparison of enrolled patients to those not enrolled. NOTE: In Chapter 4 where actual data will

136
appear, each of the analysis presented below will be performed for each of the two guidelines. For the purpose of this Chapter, the issue will be discussed only once for brevity purposes.

**Eligible versus Not Eligible**

As mentioned previously, all patients who had orders written for drug therapies covered by either of the two respective guidelines were eligible to be enrolled. Patients were not eligible for enrollment if they were not located in University Hospitals (Rhodes or Doan Hall). Thus, patients located in Dodd Hall (a rehabilitation hospital), East Hospital (another campus), and James Cancer Hospital and Research Center were not eligible for enrollment into the study. Non-eligibility was due to the Investigational Review Board processes that are not the same for each of these locations. Similarly, patients located on the 8 West Doan Unit (prisoner unit) and the 6 West Doan Unit (labor and delivery) was not eligible for enrollment due to Investigational Review Board considerations. Any patient less than 18 years old was not eligible for study inclusion due to Investigational Review Board considerations. These eligibility criteria were consistent for each of the two guidelines examined by this study. No data was available for patients not eligible for enrollment. Thus, no examination of the two groups is possible.
Enrolled Patients versus Not Enrolled Patients

For all eligible patients for enrollment into the study, the actual number enrolled will be reported. In addition, the reason for patients not enrolled into the study will be listed. Possible reasons for not enrolled include: not able due to physical limitations, patient refusal to participate or patient was not contacted. Many patients were not contacted for enrollment due to the inability of the researcher to locate the patient. Many patients were not in their hospital room when an attempt at contact was made or there were health care practitioners in the patient room or visitors were present. If the patient remained inaccessible, an attempt to contact them the next day was made. In the time lag, many patients were discharged from the hospital or expired before they were enrolled. Therefore, many were not selected due to logistical problems.

Non Response Analysis

Demographic and economic comparisons of the enrolled versus the not enrolled were compared for each respective guideline. Demographic variables compared include: admission source, discharge disposition (in-hospital mortality), and admission service. Economic variables compared include total cost of treatment and length of stay. Each of the variables were compared using Student t-tests for continuous variables and Chi-square tests for categorical variables. These statistical techniques are discussed in detail later in this Chapter.
Lost To Follow-Up

For the second health status instrument, lost to follow-up is presented. In retrospective, observational studies, lost to follow-up may be a significant source for selection bias. For each of the two guidelines, reasons for lost to follow-up were provided. In addition, comparisons of selected demographic and economic variables listed above (see non-response analysis section above) will be made for follow-up versus lost to follow-up groups. These analyses determined if there was a characteristic common to the lost to follow-up patients which were of concern to the validity of the results.

Item Non Response

Item non-response is not a significant concern for this study. The health status instruments at baseline and follow-up require complete SF-12 surveys. For item non-response on the health status instruments, the entire patient response will not be utilized in the analysis. This is the current recommendation of the developers of the SF-12 instrument. For all other data, item non-response will be treated as missing and the patient response will be included in the rest of the analysis.
Response Rate

For each of the two respective guidelines a response rate will be calculated. The numerator will consist of all patients who agreed to participate and were enrolled in the study. The denominator will consist of all patients eligible for enrollment in the study. The response rates will thus give the percentage of enrollees over those eligible for enrollment in the study.

INSTRUMENT DEVELOPMENT

Instruments used for the measurement of study variables were taken from the literature with input from patients, College of Pharmacy faculty and graduate students and practicing physicians. This section will review the data collection tool (for chart review) and the patient survey instruments (for health status and utilization measurement). Similar data was collected for the evaluation of each of the two guidelines. Thus, the discussion will focus on the data collected and source. After this discussion, the evaluation of the properties of the tools will be presented.

Demographics

Demographic data was collected from three sources. First, most of the data was collected from patient chart reviews. This data includes the admission service, age, gender, race, primary and secondary insurance type and marital status. Second, admission source was obtained from the OSUMC financial database. Third, education
level and employment status prior to admission were obtained on the baseline patient survey instrument. The demographic data was collected to report general descriptive characteristics of the sample. In addition, the demographic data that are significantly different between groups were used to account for or statistically controlled for the research questions. Since the study had a relatively short duration, no demographic data was collected in the follow-up survey instrument. It was felt that all of the demographic variables were stable over the 30-day study period.

Clinical History

Clinical history information was collected from patient charts, the baseline survey instrument and the OSUMC financial database. Data on ICD-9 codes, comorbid conditions (diabetes, renal disease, AIDS, cancer, organ transplantation, cardiovascular disease and pulmonary disease), antibiotic use on admission, fever on admission, elevated white blood cell count on admission, pulse on admission were all collected via patient chart review. Days of present illness, previous hospital, emergency department, and physician office utilization and baseline health status measurement were collected on the baseline survey instrument. The admission source variable was collected from the OSUMC financial database.

Clinical history variables were used to report general descriptive characteristics of the sample. The clinical history variables were used to account for differences found in the research questions or were statistically controlled.
Guideline Compliance or Process Variables

All variables used to classify a patient as having an order in compliance or not in compliance with practice guideline recommendations were collected via patient chart review. The data was not used for direct examination of any of the research questions. However, the data defines the groups of analyses for all of the research questions.

Economic Outcomes

The economic data were collected from two sources. First, total hospital cost, length of stay (LOS), drug therapy cost and laboratory cost were collected from the OSUMC financial database. Second, days of lost activity, hospital readmission, emergency department, physician and antimicrobial post discharge utilization rates were collected via the follow-up patient survey instrument. The economic data collected was utilized for research questions one and four.

Clinical Outcomes

Two clinical outcome variables were used in the study. The first is mortality that was collected via the OSUMC financial database. It should be noted that only in-hospital patient deaths were reported in this study. The second clinical variable was antimicrobial therapy success as defined in Chapter 1. The data used to classify a therapy as a success or failure was collected from the patient chart. This data includes
fever, WBC and pulse data, additional antimicrobial therapy, and mortality. Clinical outcome data was used in the analyses of research questions two and four.

Health Status

The SF-12 health status instrument data was collected on the baseline and follow-up patient survey instruments. The scoring of the instrument was discussed previously. The SF-12 data was collected for use in research question three.

Cost Effectiveness Analysis (CEA)

Total hospital costs as collected from the OSUMC financial database and antimicrobial success rate as collected from the patient chart were used in the CEA. This data was used for research question four.

INSTRUMENT AND TOOL EVALUATION PROCEDURES

The following discussion refers to the initial testing of the baseline patient survey that collected data on health status, recent clinical history and health care utilization. Since the follow-up survey was the same in content, it was not tested for psychometric properties. The chart review form will be discussed under the reliability section since this will be the only applicable section for the form.
Panel of Experts

Pharmacy faculty and graduate students in the College of Pharmacy and practicing physicians formed a panel of experts and reviewed the patient chart review form and the patient health status measurement instruments for content validity.

Field Test

A preliminary survey was field tested in a small group of hospitalized patients who currently had orders for vancomycin (n = 13). The purpose of the field test was to examine the wording of all items included in the survey. The group also served to test the content validity of the survey. Subjects for the field test were identified via use of the OSUMC Department of Pharmacy computer system.

No patient specific data were collected from the field test sample. Also, no incentives were given to patients to participate in the field test. The survey instrument was administered according to the procedures presented earlier in this Chapter.

Based on comments from the field test, minor wording changes were made to facilitate understanding. Most of the wording alterations were made to the utilization questions. No changes were made to the health status items of the survey.

Pilot Study

No pilot study was conducted for the purpose of this study. The purpose of a pilot study is to make an initial examination of the psychometric properties of the health
status instrument and to identify confusing or poorly worded items. Development of the SF-12 instrument and testing of the psychometric properties are reported elsewhere. The psychometric properties appear to be sound as used in past studies but will be examined in the study population since this will be a new use.

INSTRUMENT VALIDITY

Validity refers to the extent to which a measure actually measures what it says it does and is not measuring anything else. Validity answers the following question: are we measuring what we think we are measuring? Three types of validity reviewed in Chapter 2 include criterion, content and construct validity.

Criterion and Criterion-Related Validity

Criterion validity or criterion-related validity is when a scale score is correlated with some other measure of the same construct under study. For criterion validity, a 'gold standard' is used as the criterion. For criterion-related validity, a measure of a similar design, which purports to measure the same construct, is used as the criterion. Neither criterion nor criterion-related validity was assessed in the study due to a lack of a 'gold standard' or a similar measure of the construct respectively.
Content Validity

Content validity assesses if the items of an instrument are representative of the content to be measured. The method used in this study was to have the panel of experts who were familiar with the content of the instrument review and comment on the instrument. The results of the field study also support the content validity of the instrument for the health status survey. The panel of experts recommended the wording of the utilization items be specifically related to a time period rather than placing the time reference in the directions to the patient. The time reference was placed in both the directions to patients and items for the final survey.

The field test supported the content validity of the health status instrument as well. The respondents had no comments based on the survey. Respondent burden was less than five minutes that indicates that most participants in the field test did not have any problem with the clarity of the items. There were no missing items in the field test results which is an indication that there was little confusion over item wording.

The SF-12 was developed based on the SF-36 as reviewed earlier in Chapter Two. This data supported the content validity of the instrument. Therefore, no data is presented on the content validity of the SF-12.

Construct Validity

Construct validity is the extent to which a measure behaves the way that the construct it purports to measure should behave with regard to established measures of
other constructs. The SF-12 will be evaluated for use in the current study population. Since each of the guidelines focused on antimicrobial therapy, the data was pooled for test validity. Known group comparison of the Physical Component Summary (PCS-12) and the Mental Component Summary (MCS-12) were not done for this study. Due to the severity of illness issues of the patient population where most of the enrolled patient had one or more comorbid conditions, known group definition was too difficult. However, patients will be identified with ICD-9 codes for mental illness for known group construct validity testing. The MCS-12 and PCS-12 were correlated for patients with a mental illness ICD-9 code and for patients without the ICD-9 codes. The MCS-12 for patients with the ICD-9 codes for mental illness should be lower than the MCS-12 for patients without the ICD-9 codes. The number of comorbidities was correlated with the PCS-12 and the MCS-12 scales. The research hypothesis tested is that as the number of comorbid conditions increase a correlated decrease in the two scales should occur.

Since there is no other measure or method used in the study, discriminate or multitrait-multimethod techniques were not used to test construct validity.

RELIABILITY/HEALTH STATUS INSTRUMENT

Measurements are reliable to the extent that they are repeatable and that any random influence, which causes different measurements of the same variable to vary, is a source of measurement error. For the purpose of testing the instrument, this study
again pooled the data from the two guidelines to test reliability. Test-retest reliability was conducted in this study. However, caution should be exercised since the patients enrolled in the study were not stable clinically. Also, internal consistency was calculated for the PCS-12 and MCS-12, respectively. This test of reliability will be conducted on all data collected in the baseline measurement. Cronbach’s $\alpha$ was used to test internal consistency and is defined by the following equation:\textsuperscript{11}

$$\alpha = \frac{k}{k-1} \frac{1 - \sum \sigma_i^2}{\sum \sigma_i^2}$$

where, $k =$ number of items
$\sigma_i^2 =$ item variance
$\sigma_t^2 =$ total variance

Equation 3.1

Cronbach’s $\alpha$ estimates the reliability based on the average correlation among items within a given score. Item variance ($\sigma_i^2$) refers to the item variances or the diagonals of a covariance matrix and total variance ($\sigma_t^2$) refers to the off diagonals plus the diagonals of a covariance matrix. For group comparisons, reliability coefficients of 0.70 or greater are generally satisfactory.\textsuperscript{10}

Another issue examined in the study concerns the ‘ceiling’ and ‘floor’ effects. The ceiling effect may be defined as scoring the highest possible score on a test that does not allow for improvement. The floor effect may be defined as scoring at the lowest score that does not allow for a decline in score. For this study, a major concern
is 'floor'effects. Since patients are very ill and in the hospital, significant floor effect scores may be a threat to the reliability of the instrument for use in this population. The rate of ceiling and floor effects were reported in this study.

RELIABILITY/CHART REVIEW FORM

The reliability of the chart review process was examined in the study. A clinical nurse practitioner reviewed approximately 8.0% of the patient (n = 10) charts for the vancomycin guideline and (n = 14) the iv to po switch guideline, respectively. The original researcher chart review results were compared to the clinical nurse practitioner chart reviews to assess agreement. In addition to agreement, the Kappa measure of agreement was used to measure interrater agreement. The Kappa statistic works by testing if the counts in the diagonal cells (where agreement exists) in an R by R table differ from those expected by chance alone. The following equation is used for the calculation of Kappa:

\[
\kappa = \frac{p_o - p_e}{1 - p_e}
\]

where, \(\kappa = \) Kappa statistic
\(p_o = \) the sum of the observed proportions in agreement
\(p_e = \) the sum of the expected proportion in the same cell

Equation 3.2

Two Kappa measure were calculated in the study. The first is agreement on the judgment if the respective drug therapy order was compliant or not compliant with the
practice guideline recommendations. The second was a measure of agreement of the judgment of success for the antimicrobial therapy. Values of Kappa greater than 0.75 indicate excellent agreement beyond chance, values between 0.40 and 0.75 indicate fair to good agreement values and values below 0.40 indicate poor agreement.²¹

RESPONSIVENESS

Responsiveness or sensitivity to change refers to the instrument's ability to detect clinically important changes over time, even if those changes are small.¹² For the purpose of the study, the standardized response mean (SRM) was calculated.¹³ The SRM was used in a previous study using the SF-12 by Hurst et al.¹⁴ SRM may be defined as the mean change in score divided by the standard deviation. The higher the value of SRM indicates greater responsiveness. SRM will be calculated using the following equation:¹³

\[
ES = \frac{(M_2 - M_1)^2}{SD_1}
\]

where, \(ES\) = effect size
\(M_1\) = baseline mean
\(M_2\) = follow-up mean
\(SD_1\) = sum of difference in standard deviation of \(M_1 - M_2\) for individuals.

Equation 3.3

Second, global ratings of patient physical function and mental function were included in the patient follow-up survey. The patient was asked if their overall health
function had deteriorated, remained stable, or improved since they were discharged from the hospital. This method is similar to the method used to test responsiveness as Guyatt et al.\textsuperscript{15} One would expect an improvement in score that across all subjects who improved between measurements. Paired t-tests will be calculated to test if these differences in scores are statistically significant.

TARGET POPULATION AND SAMPLING

In this section of Chapter 3, the target population is defined. Then the sampling procedure is explained in detail. Finally, the necessary sample size is calculated based on the assumptions provided.

Target Population

The target population for the study was all patients treated at OSUMC who required one of the drug therapies covered by the respective practice guideline. For the intravenous (iv) too po guideline the target population included all patients who had an intravenous antimicrobial drug therapy order written for intravenous ciprofloxacin, azithromycin, clindamycin, metronidazole, or ampicillin/sulbactam (see Table 3.1). For the vancomycin guideline the target population was all OSUMC patients who had orders written for intravenous (IV) vancomycin.
The inclusion criteria for the study consisted of patients who: 1) had a written order for a drug therapy per the guideline, 2) signed a patient informed consent to participate in the study, and 3) were capable of completing the patient survey.

Thus, the only exclusion criteria for the study were patients physically unable to complete the necessary baseline survey. This fact may be a limitation as discussed later in this Chapter.

**Sampling Procedures/Data Collection**

The sampling procedure for each of the two analysis included enrolling patients who met the inclusion criteria (see target population section above). This sampling procedure may be best described as a non-probability, convenience sample. Potential patients were identified through the use of the computer system of the Department of Pharmacy at OSUMC. A daily list of potential patients was printed and contained the name of the medication ordered, the patient's name, and the patient's hospital room number. To maintain patient confidentiality, these lists were destroyed at the end of each day. Thus, the researcher worked from a new list each day during the course of the study.

The researcher then personally contacted each potential patient. During the initial contact, the patient was given a copy of the informed consent form (see Appendix E). The patient was instructed to read the informed consent form in the presence of the researcher. After the patient was finished reading the information, the researcher asked
if the patient had any questions or concerns regarding the study purpose or procedures. The patient was then asked if they were willing to participate. If a patient declined to participate, the researcher thanked the patient and recorded the fact that the patient did not wish to participate. This information was used to ensure a patient would not be contacted a second time regarding study participation.

If the patient agreed to participate in the study, they were instructed to sign the informed consent form in the designated areas. The researcher also signed and dated the consent form in the presence of the patient. The patient was given the initial baseline health status instrument. The instrument had a cover that had the seal of The Ohio State University printed on it. After the patient completed the in-hospital baseline survey, the researcher emphasized that the patient would receive a second similar survey in the mail approximately thirty days after they were discharged from the hospital. Participating patients were told that the second survey was critical to the research and that a self-addressed, pre-paid envelope would be provided. A cover letter printed on official Ohio State University, College of Pharmacy letterhead was also included (see Appendix J).

Two weeks after mailing the follow-up survey, participating patients were contacted by telephone by the researcher. The patient was reminded of the importance of the follow-up survey or the survey was conducted by telephone.

The same researcher conducted all patient contact and surveys. The researcher followed a structured routine. A few minutes of introduction and patient visiting were
done at the beginning and end of each contact. This was done to put the patient at ease.

No patient incentives were given to patients.

The data collected from patient charts was collected retrospectively. All chart data was collected based upon using structured data collection forms (see Appendix G).

Sample Size

The power of a statistical test is the probability that it will yield statistically significant results.\(^\text{16}\) Unfortunately statistical power is often misunderstood and is often not discussed when it has a clear influence on the interpretation of study results. Any result taken from a sample drawn from a population will at best approximate the characteristic of the population. The investigator selects a significance criterion prior to the study. For the purpose of this study a significance criterion of 0.05 will be used (e.g., \(\alpha = 0.05\)). This value was selected to reject the null hypotheses stated for the study at a 0.05 significance level. Null hypotheses will be 'rejected' or 'fail to reject' based upon the significance criterion. A second issue involved in power is the assumption that all tests were two-sided or two-tailed. Since there is no \textit{a priori} evidence of directional relationships among any of the hypothesized variables, two-sided assumptions were appropriate. The power of a statistical test of null hypotheses is the probability that it would lead to the rejection of the null hypotheses or the probability that it would result in the conclusion that the phenomenon exists.\(^\text{16}\)
The power of a statistical test depends on three parameters: 1) the significance criterion, 2) the reliability of the sample results, and 3) the 'effect size' or the degree to which the phenomenon exists. The first parameter has been previously discussed. However, one additional point should be emphasized. Power is inversely related to the significance criterion or $\alpha$. The smaller the $\alpha$ selected for a study, the lower the probability of making a Type I error (e.g., rate of rejecting a true null hypothesis). At the same time, this reduces the power of the statistical test. The compliment of power, $(1 - \text{power})$ is termed $\beta$ or Type II error (e.g., rate of failing to reject a false null hypothesis). Thus a proper balance must be reached between the researchers' willingness to make a Type I error versus a Type II error. Based upon the assumptions discussed in detail below, this project has set $\alpha = 0.05$ and power = 0.80 ($\beta = 0.20$). Under these assumptions, the researcher has implicitly set the relative seriousness of Type I to Type II error at 4 to 1. Thus, a mistaken rejection of the null hypotheses is considered four times as serious as mistaken acceptance of the null hypotheses.

For the second parameter, reliability of the sample results, little discussion is needed. The reliability of the sample results are always a function of the sample size. As a sample gets larger, variation or closeness to the population value improves. The larger the sample size, the smaller the error and the greater the reliability or precision of the results. Thus, increasing the sample size will always increase the power of a test. Other random measurement error can also effect the power of a test. Thus, a reduction in measurement error will result in increased power.
The third and final parameter in power analysis is the 'effect size'.\textsuperscript{16} Effect size may be defined as 'the degree to which the phenomenon is present in the population' or 'the degree to which the null hypothesis is false'.\textsuperscript{16} Effect size is the most important determinant of power or required sample size.\textsuperscript{16} The larger the effect size posited in a study, other things (significance criterion, sample size) being equal, the greater the power of the test. Alternatively and more importantly for the purpose of this discussion, the larger the effect size posited, other things (significance criterion, desired power) being equal, the smaller the sample size necessary to detect it.

For the purpose of this study, \( n \) or the sample per group will be a function of three parameters (effect size, significance criterion and power). When any investigator anticipates a certain effect size, sets a significance criterion or \( \alpha \) and specifies the amount of power desired, the sample size necessary to meet the specifications can be determined. The current study will use the main clinical outcome variable as the phenomenon to calculate sample size. This variable is the rate or proportion of patients who have successful antimicrobial therapy as compared between two independent populations (e.g., patients with orders compliant with practice guideline recommendations and patients with orders not compliant with practice guideline recommendations). Given a significance criterion of 0.05, the desired power of 0.80 and an effect size of 0.15 (e.g., a medium effect as defined by Cohen et al.) and using the power tables presented by Cohen et al. a sample size of 63 per group is required in the study.\textsuperscript{16} The effect size represents the difference in proportions between the two
groups which was deemed clinically significant.\textsuperscript{16} Reported failure rates in the literature for antimicrobial therapy in the general population has been estimated at 10 to 40\%.\textsuperscript{17} Similarly, for antimicrobial failure rate when switching from intravenous to oral has been reported to range between 5 and 16\%.\textsuperscript{18,19} Thus, the effect size assumes that following the guideline will have a slightly improved rate of success of 15.0\%.

Using the same assumptions as stated above, the sample size was calculated using SamplePower\textsuperscript{TM} 1.00, SPSS Inc., Chicago, IL.\textsuperscript{20} A sample size of 65 per group was calculated using this program.

STATISTICAL ANALYSIS

A review of the statistical analyses utilized to answer the research questions posed in the study is provided in this section of the Chapter. Detailed discussion of the statistical analysis will not be repeated for each research hypothesis. Rather, when a new statistical technique is used it will be discussed in detail. Later hypothesis will refer to the test already covered in the section. Compliance is defined by following the recommendations. This cohort of patients treated by physicians who were compliant with the recommendations will be referred to as the compliant group of patients (CGP). Non-compliance is defined as not following guideline recommendations. This cohort of patients will be referred to as the non-compliant group of patients (NCGP).
Each guideline will be analyzed separately using the same set of research questions and statistical analysis. The following reviews the two statistical techniques encountered in research: the majority of the research questions below.

**Independent Samples t-test**

The first test when conducting an independent samples t-test is to determine if the variance of the two groups were equal or unequal. This was done through the use of Levene’s Test for Equality of Variance. This tests if the dispersion of the variable for the groups differed. If the variances were equal, then the pooled-variance t-test was the appropriate analysis. If the variances were not equal, then the separate-variances t-test was the appropriate test statistic. The formula for the two different types of t-tests are given below:

\[
Pooled\text{-}variance\ t\text{-}test = \frac{x_1 - x_2}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}
\]

where, \(s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\)

\[
Separate\text{-}variance\ t\text{-}test = \frac{x_1 - x_2}{\sqrt{(s_1^2/n_1 + s_2^2/n_2)}}
\]

where, \(X_i = \text{mean of population I}\)
\(n_i = \text{number of observations in group I}\)
\(s_i^2 = \text{sample variance in group I}\)
In the computer software used to support this study, the data for both the pooled and separate t-tests are given. If the variances are determined to be equal in the Levene test, then pooled t-test results will be reported. If the variances are unequal, then separate variance t-tests will be interpreted.

Chi – Square Test of Independence

This is a test statistic was used as a measure of association between two variables or to test whether the variables are independent of one another. In a R by C table with the Chi-square test of independence a two way frequency table is formed with R rows and C columns. It should be noted that in a 2 by 2 column table, Fisher's exact test will be reported. For a square table such as R rows and R columns (R by R) the Kappa measure of agreement as discussed earlier will be reported. The interpretation of the result for all of these measures are similar.

For tables with R by C format, the Pearson chi-square statistic is utilized to test the independence of table rows and columns. For example, one might want to test the premise that compliance and gender are independent of each other. Thus, the Pearson chi-square tests if the row (compliance) and column (gender) are independent. Two variables are independent if the probability that a case falls in a specific cell is the product of its marginal probabilities. The example below will be used for illustrative purposes:
Table 3.2: Illustrative Chi-Square Table

Using a probability that a subject is in row D or (K/O) and the probability that a subject is in column A or (H/O), the probability for a case falling in the upper left cell (a) is: \( \frac{K \times H}{O} \). This probability is used to estimate the number of cases expected in each cell in the table. The expected count is then compared with the observed count.

To compute the expected number of cases, multiply the probability by the total sample size. This result (cases expected for cell a) is the row total (K) x column total (H)/total sample size (O).

The difference between the observed count and the expected count is calculated. This process is repeated for each cell and the Chi-square is given by the following equation:\(^{11}\)
\[ \chi^2 = \Sigma_i \Sigma_j \frac{(o_{ij} - e_{ij})}{e_{ij}} \]

where, \( \Sigma_i \) = summation across all rows
\( \Sigma_j \) = summation across all columns
\( o_{ij} \) = observed count in row I, column j
\( e_{ij} \) = expected count in row I, column j

Equation 3.5

When a chi-square is large, the null hypothesis of independence is rejected. To define large, a probability of the result less than 0.05 (the significance level) will result in the null hypothesis being rejected. Two rules will be observed for use of the Pearson Chi-square. First, no cell may have an expected value less than one. Second, no more than 20% of the cells can have expected values less than 5. In the computer statistics package this information is automatically reported. When either of these scenarios occurs, data collapsing or deletion may be appropriate.

The Chi-square statistic can only tell the researcher if there is an association between two variables. It does not tell the researcher how the variables are related or how strong the relationship is.

For each of the hypotheses measures of symmetry were reported for each Chi-square test. Detailed discussion of these measures is beyond the scope of this discussion.
Research Question One

Are there economic outcomes for the CGP significantly different from the NCGP? One of the main tenants of the use of practice guidelines for drug products is utilization rates (i.e., cost reduction). For all cost variables, estimated cost will be utilized. For the purposes of the study, OSUMC cost estimates will be used and are calculated by the use of cost-to-charge ratios. Currently, OSUMC does not utilize a cost accounting method. Therefore actual cost incorporating overhead costs directly into the measurement were not available. However, overhead costs were estimated at the operational level throughout the hospital and incorporated into departmental cost-to-charge ratios. The term cost used throughout the study refers to estimated cost. It was also felt that since the study is being conducted at the same institution, cost estimates will be appropriate since they are consistent between patients enrolled in the study.

Some of the variables listed below refer to resource utilization only. They were simple measures of resources consumed. For the statistical analysis in this section, if significant demographic or clinical history variables were found, then these variables will be used to account for differences or were statistically controlled for. If no differences were found, then independent samples t-tests will be performed.

1.1 Do mean total hospital costs for the CGP differ from the mean of NCGP?
1.2 Do mean total hospital LOS for the CGP differ from the mean of the NCGP?
1.3 Do mean drug therapy costs for the CGP differ from the mean of the NCGP?
1.4 Do mean laboratory costs for the CGP differ from the mean of the NCGP?
1.5 Do mean number of days of lost activity in the first 30 days after discharge for the CPG differ from the mean of the NCGP?

1.6 Do mean number of hospital readmissions for infection in the first 30 days after discharge for the CPG differ from the mean of the NCGP?

1.7 Do mean number of ED visits for infection in the first 30 days after discharge for the CPG differ from the mean of the NCGP?

1.8 Do mean number of physician office visits for infection in the first 30 days after discharge for the CPG differ from the mean of the NCGP?

1.9 Is the mean number of patients receiving new antimicrobial therapy in the 30 days after discharge in the CPG different from the mean in the NCGP?

**Analysis of Covariance**

The dependent variable (a continuous variable such as total cost in Research Hypothesis 1.1) is what will be examined in this analysis. The issue is that if there were significant differences between the groups at baseline with respect to baseline demographic or clinical history variables, then the differences must be accounted for in the analysis or used to explain variance. For example, for this discussion mean age between the groups will assumed to be different. Since there was variation in the baseline groups, the baseline value (age) is entered as a covariate in the model in order to determine whether order compliance makes a difference in total cost.
In order to be able to compare costs for the two groups without having to condition on a particular value of the covariate (age in our fictional example), the researcher must be able to assure that the regression of cost on age is the same for each of the two groups (compliant versus non-compliant orders). This assumption of equality (homogeneity) of regression slopes can be tested by fitting a model containing main effects of group and age as well as the group*age interaction. This interaction term provides the test of the null hypotheses of equal slopes. If the homogeneity of regressions assumption is not rejected, the researcher may proceed to estimate the effects of group on total cost given baseline age differences between the groups.

The interpretation of analysis of covariance is not a major purpose for this research. F values and probabilities were used to interpret main effects in the ANCOVA models if and when they were used.

Research Question Two

Do patients treated in the guideline CGP have better clinical outcomes than the patients treated in the NCGP? Compliance with the guideline may produce clinically positive outcomes, negative outcomes, or have no effect on clinical outcomes. The main effectiveness measure used in the study was antimicrobial therapy failure rate. For each guideline, the defined failure rate represented the number of failures divided by the number of patients in the study group. For the iv to po guideline, failure is defined as the
switch from iv to po therapy and then a subsequent switch back to an iv therapy or
death. For the vancomycin guideline, failure is defined as the patient having continued
symptomology of infection (fever, increased WBC and/or high pulse rate) or the
addition of another antibiotic or death. The following research hypotheses are designed
to examine research question 2 and were tested using Chi-Square tests of independence.

2.1 Does the in-hospital mortality rate for the CGP differ from the rate in the
NCGP? (Chi – Square Test of Independence)

2.2 Does the drug therapy failure rate for the CGP differ from the rate for the
NCGP?

Research Question 3:

The SF-12 is a multipurpose short-form (SF) generic measure of health status. The
SF-12 measures eight domains commonly represented in many health status surveys
and include: physical functioning, role limitations due to physical health problems,
bodily pain, general health, vitality (energy/fatigue), social functioning, role limitations
due to emotional problems, and mental health (psychological distress and psychological
well being). The SF-12 was developed due to concerns over respondent burden. The
SF-12 generally takes 2 to 3 minutes to complete. The main purpose for the use of the
instrument in this study is to capture a measure of health status and yet hold respondent
burden low.
Although brief, the SF-12 accounts for more than 90% of variance as measured by the SF-36 instrument\(^7\). The SF-36 is a longer instrument from which the SF-36 was developed. The scores for the SF-12 correlate highly with the scores for the SF-36. Ware et al. has compared the results of the SF-12 and the SF-36 in patients with four different medical conditions that differ from each other in terms of severity\(^7\). For each of the four conditions, the SF-12 produced slightly less valid results, but the statistical conclusions for the two instruments were identical\(^7\). Similar reliability and validity tests were conducted for patients with sixteen co-morbid conditions, patients with five different acute conditions and various age groups of healthy people. As compared to the SF-36, the SF-12 produced the same statistical conclusions.

Do patients in the CGP report better health status outcomes than those in the NCGP? A significant feature of this study is the ability to examine the effect of physician compliance with guideline recommendations on patient health status. All of the health status hypotheses will be analyzed for the two summary scores of the SF-12 (i.e., the Physical Component Summary and the Mental Component Summary). The following research hypotheses will be used to examine this research question:

3.1 Do mean health status scores for the CGP differ from the mean for NCGP at 30 days post discharge?
3.1.1 Does the mean SF-12 Physical Component Summary score for the CPG differ from the mean for NCGP?

3.1.2 Does mean SF-12 Mental Component Summary score for the CGP differ from the mean for NCGP?

3.2 Is the mean difference score (follow-up minus baseline score) for the CGP different from the mean difference score for the NCGP?

3.2.1 Is the mean difference SF-12 Physical Component Summary score for the CPG different from the mean for NCGP?

3.2.2 Is the mean difference SF-12 Mental Component Summary score for the CGP different from the mean for NCGP?

**Research Question Four**

In this study, estimated total hospital costs will be used as the cost portion of the cost effectiveness ratio. The success rate for each guideline (as defined previously) will be used for the effectiveness portion of the ratio. The reader is referred to Chapter 2 for a complete discussion of CEA procedures to be followed in this study.

4.1 Is the cost effectiveness for the CGP different from the cost effectiveness for the NCGP?
DESIGN LIMITATIONS

Campbell et al. presents several threats to internal and external validity. Internal validity concerns whether the experimental treatments make a difference in this experimental instance. Internal validity is concerned with extraneous variable effects that may confound with the effect of the experimental stimulus. This study has the following threats to internal validity: history, testing, statistical regression, selection bias, experimental mortality, and selection-maturation interaction.

**History**

History effects occur when specific events occur between the first and second measurement in addition to the experimental variable. Since many of these patients were extremely ill, many factors may have occurred in their health over thirty days. These events may have influenced the health status measurement in the study.

**Testing**

Testing refers to the effects of taking a test upon the scores of taking a second test. In this study this threat may have been relevant to the health status measurement. However, 30 days was allowed to elapse between testing which is usually a sufficient wasing-out period.
Statistical Regression

Regression is a concern when groups have been selected on the basis of belonging to an extreme group.¹ This study examines extremely ill patients in a hospital setting. Thus, regression is a concern for the health status score. Another related issue is the fact that the ‘sickest’ of the ill patients were not able to participate in the study. Thus, regression is a concern.

Selection

Selection refers to the bias resulting from differential selection of respondents for comparison groups.¹ There has been significant discussion and attempts to identify sources for this threat. This is the most significant threat to an observational study design.

Selection-Maturation Interaction

This threat is where the selection interacts to be confounded as an effect of the study.¹ This threat is not a significant threat to the study, but needs to be kept in mind for the selection issue mentioned previously.

External Validity

External validity is concerned with the generalizability.¹ What populations, settings, treatment variables, and measurement variables can this effect be generalized.
The Reactive or Interaction Effect of Testing

This is not a significant threat for this study. No pretest was done in the study. For the field test, a separate sample of patient was utilized. Thus, there should no have been any sensitization of the study group.

Interaction of Selection Biases and the Experimental Variable

This as mentioned earlier in the study is a significant threat to external validity. It is possible that selection bias might influence the results obtained. This scenario would strongly limit the generalizability of the study results.

Reactive Effects of the Experimental Arrangements

This threat is not existent in this study. Since the study was conducted in a real setting, results should be generalizable to this setting.
REFERENCES FOR CHAPTER 3


7. Ware JE, Kosinski M and Keller SD. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales. Boston, MA;The Health Institute, New England Medical Center;1995.


11. SPSS® Base 9.0 Applications Guide. Chicago, IL;SPSS, Inc.;1999.


CHAPTER 4

RESULTS

This chapter is divided into five major sections. First, guideline evaluation is discussed. This section describes the results used to examine the quality of the practice guidelines included in this study. Then, evidence of reliability and validity for the study health status instrument (SF-12) is provided. This analysis was conducted using pooled data from two guidelines. The third section divides the results for each of the two guidelines analyzed. Thus, for each individual guideline description of the sample of patients enrolled in the study is presented through the reporting of response rate and non-response. Fourth, the study sample is described. Basic descriptive results and clinical history comparisons are provided in this section. Finally, the results of the statistical analyses corresponding to the four research questions are presented.
GUIDELINE EVALUATION

Two physicians were asked to rate the two practice guidelines using the criteria and methodology used by Shaneyfelt et al.¹ Disagreements between rater were resolved through open discussion until agreement was reached. Table 4.1 presents the results of the evaluation. Overall, for the iv to po guideline, 60.0 percent of the development standards were met (15 of 25). For the vancomycin guideline, 64.0 percent of the development standards were met (16 of 25).

For methodological standards (e.g. standards 1 to 10) on guideline development and format, both guidelines were rated as meeting 8 of 10 standards. For methodological standards on evidence identification and summary (e.g. standards 11 to 20), both guidelines were rated to meet 5 of 10 standards. Finally, for methodological standards on the formulation of recommendations (e.g. standard 21 to 25) the iv to po guideline was rated to meet 2 of 5 while the vancomycin guideline met 3 of 5. The results indicate the two guidelines meet adequate criteria as found in the literature.¹
### Standards of Guideline Development

<table>
<thead>
<tr>
<th>Standards of Guideline Development</th>
<th>Converting Intravenous to Oral Antimicrobials</th>
<th>Criteria for Use of Vancomycin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Purpose of the guideline is specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2 Rationale and importance of the guideline are explained</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>3 The participants in the guideline development process and their areas of expertise are specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>4 Targeted health problem or technology is clearly defined</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>5 Targeted patient population is specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>6 Intended audience or users of the guideline are specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>7 The principal preventive, diagnostic, or therapeutic options available to clinicians and patients are specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>8 The health outcomes are specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>9 The method by which the guideline underwent external review is specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>10 An expiration date or date of scheduled review is specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>11 Method of identifying scientific evidence is specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>12 Time period from which evidence is reviewed is specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>13 The evidence used is identified by citation and reference</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>14 Method of data extraction is specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>15 Method for grading or classifying the scientific evidence is specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>16 Formal methods of combining evidence or expert opinion are used and described</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>17 Benefits and harms of specific health practices are specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>18 Benefits and harms are quantified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>19 The effect on health care costs from specific health practices is specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>20 Costs are quantified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>21 The role of value judgments used by the guideline developers in making recommendations is discussed</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>22 The role of patient preferences is discussed</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>23 Recommendations are specific and apply to the stated goals of the guideline</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>24 Recommendations are graded according to the strength of the evidence</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>25 Flexibility in the recommendations is specified</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>TOTAL STANDARDS ENDORSED</strong></td>
<td><strong>15/25</strong></td>
<td><strong>16/25</strong></td>
</tr>
</tbody>
</table>

Table 4.1 OSUMC Guideline Development Standards
PSYCHOMETRIC PROPERTIES OF INSTRUMENTS

This section reviews the psychometric properties tested in the study. For psychometric testing the results obtained for the two practice guidelines were pooled to increase the power of the tests. Additionally, since each guideline dealt with antimicrobial drug therapy, it was felt appropriate to do combine the data.

Validity

Tables 4.2 and 4.3 present the correlations of PCS-12 and MCS-12 with number of comorbid conditions. For each of these analyses, the correlations were negative and statistically significant. Correlation was $-0.139 \ (p = 0.031)$ for the PCS-12 while it was $-2.43 \ (p < 0.001)$ for the MCS-12 with number of comorbidities. In other words, as the number of comorbid conditions increase, the score of the PCS-12 and MCS-12 decrease respectively. Table 4.4 presents mean comparison on the MCS-12 for patients with a mental illness ICD-9 code to those patients without an ICD-9 code. The results were significantly different ($t = -2.323, \ p = 0.023$). The results obtained support the validity of the SF-12 in the study population.
### Table 4.2 Validity Correlations – PCS-12 By Number of Comorbid Conditions

<table>
<thead>
<tr>
<th>Variable</th>
<th>PCS-12</th>
<th>Number of Comorbid Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCS-12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>1.000</td>
<td>-.139</td>
</tr>
<tr>
<td>Significance*</td>
<td>.031</td>
<td></td>
</tr>
<tr>
<td>Sum of Squares and Cross Products</td>
<td>18053.5</td>
<td>-394.2</td>
</tr>
<tr>
<td>Covariance</td>
<td>74.91</td>
<td>-1.64</td>
</tr>
</tbody>
</table>

### Table 4.3 Validity Correlations – MCS-12 By Number of Comorbid Conditions

<table>
<thead>
<tr>
<th>Variable</th>
<th>MCS-12</th>
<th>Number of Comorbid Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCS-12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>1.000</td>
<td>-2.43</td>
</tr>
<tr>
<td>Significance*</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>Sum of Squares and Cross Products</td>
<td>46368.1</td>
<td>-1104.9</td>
</tr>
<tr>
<td>Covariance</td>
<td>192.40</td>
<td>-4.60</td>
</tr>
</tbody>
</table>

### Table 4.4 Comparison of Mean MCS-12 Scores By Mental Illness ICD-9 Code

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>t</th>
<th>df</th>
<th>Sig.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with mental illness ICD-9 code</td>
<td>89</td>
<td>35.98</td>
<td>15.86</td>
<td>-2.323</td>
<td>304</td>
<td>0.023</td>
</tr>
<tr>
<td>Patients without mental illness ICD-9 code</td>
<td>216</td>
<td>41.23</td>
<td>12.93</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a - statistical significant p < .05
Reliability – Health Status

The reliability coefficients were calculated for each of the SF-12 subscales for baseline and follow-up measurements. Cronbach’s $\alpha$ tests the internal consistency or homogeneity of items for a scale. For the baseline measurement, the PCS-12 Cronbach $\alpha$ was 0.6411 while for the MCS-12 it was 0.7053. For the follow-up measurement, the PCS-12 Cronbach $\alpha$ was 0.6888 while for the MCS-12 it was 0.6844. These results were slightly low according to the standards presented by Nunnally et al. These authors recommend that reliability coefficients be 0.70 for tools used in group comparison. However, for the purposes of the current study these reliability coefficients were adequate.

The rate of ‘ceiling’ and ‘floor’ effect was examined. It was expected that due to high acuity level of patients, a higher level of floor effect would be found for PCS-12 and MCS-12 at baseline as compared to follow-up measurement. The results presented in Table 4.5 show that results were in the expected direction. The ‘floor’ effect for both the PCS-12 and MCS-12 were reduced from baseline from 15.4% and 7.5% to 4.6% and 2.6%, respectively, at the follow-up measurement. The follow-up values for both ‘floor’ effect and ‘ceiling’ effect were below the normative value of 15.0% suggested by McHorney et al.
### Table 4.5 ‘Ceiling’ and ‘Floor’ Effects of the SF-12 Subscales

<table>
<thead>
<tr>
<th></th>
<th>’Floor’ Effect n (%)</th>
<th>’Ceiling’ Effect n(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCS-12 (baseline)</td>
<td>47 (15.4)</td>
<td>1 (0.3)</td>
</tr>
<tr>
<td>MCS-12 (baseline)</td>
<td>23 (7.5)</td>
<td>11 (3.6)</td>
</tr>
<tr>
<td>PCS-12 (follow-up)</td>
<td>14 (4.6)</td>
<td>5 (1.6)</td>
</tr>
<tr>
<td>MCS-12 (follow-up)</td>
<td>8 (2.6)</td>
<td>8 (2.6)</td>
</tr>
</tbody>
</table>

#### Reliability – Chart Review

Two reviewers independently reviewed approximately 8% of all charts (n=25). Overall, agreement (Table 4.6) was found to be high between the chart reviewers on two main outcome variables – order compliance and success. Order compliance was defined as following the recommendations of each guideline (see Appendices A and B). Antimicrobial success was defined as having no signs and symptoms of infection, remaining alive and having no additional antimicrobial therapy added to a patient’s regimen. Kappa statistic for agreement was also calculated for the two outcome variables. Kappa accounts for the probability of agreement by chance alone. From Table 4.6, the reliability and agreement between reviewers on chart reviews were 0.89 and 0.89 for vancomycin charts and 0.93 and 0.92 for iv to po charts. These values are higher than normative values of 0.70 proposed by Nunnally et al.
Table 4.6 Chart Review Agreement (Reliability)

<table>
<thead>
<tr>
<th>% Agreement (compliance)</th>
<th>0.900</th>
<th>0.928</th>
</tr>
</thead>
<tbody>
<tr>
<td>κ (compliance)</td>
<td>0.890</td>
<td>0.925</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>% Agreement (success)</th>
<th>0.900</th>
<th>0.928</th>
</tr>
</thead>
<tbody>
<tr>
<td>κ (success)</td>
<td>0.890</td>
<td>0.922</td>
</tr>
</tbody>
</table>

a - 10 charts reviewed by two reviewers  
b - 14 charts reviewed by two reviewers

**Responsiveness**

Responsiveness or the ability of an instrument to detect change was measured as the standardized response mean (SRM). For the formula and rationale which underlie this calculation please see Chapter 2. The SRM for the PCS-12 in this study was calculated as 0.75 while for the MCS-12 the calculation was 0.27. As compared to the study by Hurst et al. the SRM for the PCS-12 is better in this study (0.52 versus 0.75). By contrast, the SRM for the MCS-12 was slightly lower in this study as compared to the Hurst study (0.31 versus 0.27). It should be remembered that the study by Hurst et al. was conducted on rheumatoid arthritis patients. Thus, the severity of illness in the population in the current study was more acute and severe as compared to the Hurst study. This may explain why higher responsiveness was obtained for the PCS-12 and similar results between the two studies occurred for the MCS-12.

A second method to examine responsiveness was similar to the methodology proposed by Guyatt et al. Each respondent was asked in the follow up survey to
globally rate if they improved, remained the same or worsened. From Table 4.7 and 4.8 it is obvious that the correct sequence is observed for raw difference scores. The patients with a global rating of improved had higher scores on the PCS-12 and MCS-12 scales, respectively. The order of raw difference scores go in order from improved (PCS-12 = 13.18 and MCS-12 = 11.41), to remain the same (PCS-12 = 5.88 and MCS-12 = 1.59), to worsened (PCS-12 = -3.90 and MCS-12 = -0.37). To account for baseline SF-12 scores, Tables 4.9 and 4.10 used analysis of covariance to address this issue. The dependent variable (PCS-12 and MCS-12 difference scores) was examined using the global rating as the fixed factor and baseline score as the covariate. Even after the effect of baseline was statistically controlled, the main effect for both the PCS-12 and MCS-12 difference score was significant. The patients who rated themselves as improved on the global rating, also improved the most on the scale difference score. Therefore, the PCS-12 and MCS-12 subscales were responsive to change.

<table>
<thead>
<tr>
<th>Global Rating</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved</td>
<td>117</td>
<td>13.1782</td>
<td>9.4019</td>
</tr>
<tr>
<td>Remained Same</td>
<td>119</td>
<td>5.8808</td>
<td>7.3402</td>
</tr>
<tr>
<td>Worsened</td>
<td>69</td>
<td>-3.8972</td>
<td>7.1038</td>
</tr>
<tr>
<td>Total</td>
<td>305</td>
<td>6.4537</td>
<td>10.3738</td>
</tr>
</tbody>
</table>

Table 4.7 PCS-12 Difference Score As Assessed by Global Ratings

<table>
<thead>
<tr>
<th>Global Rating</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved</td>
<td>117</td>
<td>8.5490</td>
<td>11.4091</td>
</tr>
<tr>
<td>Remained Same</td>
<td>119</td>
<td>1.5911</td>
<td>8.1002</td>
</tr>
<tr>
<td>Worsened</td>
<td>69</td>
<td>-0.3662</td>
<td>7.3802</td>
</tr>
<tr>
<td>Total</td>
<td>305</td>
<td>3.8069</td>
<td>10.0779</td>
</tr>
</tbody>
</table>

Table 4.8 MCS-12 Difference Score As Assessed by Global Ratings
Table 4.9    PCS-12 Difference Score Tests of Between Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>11897.259</td>
<td>3</td>
<td>3965.753</td>
<td>71.787</td>
<td>0.000</td>
<td>.494</td>
</tr>
<tr>
<td>Intercept</td>
<td>4502.870</td>
<td>1</td>
<td>4502.870</td>
<td>81.309</td>
<td>0.000</td>
<td>.269</td>
</tr>
<tr>
<td>PCS-12(baseline)</td>
<td>2515.428</td>
<td>1</td>
<td>2515.428</td>
<td>45.533</td>
<td>0.000</td>
<td>.171</td>
</tr>
<tr>
<td>Global Rating</td>
<td>8366.052</td>
<td>2</td>
<td>4183.026</td>
<td>75.720</td>
<td>0.000</td>
<td>.407</td>
</tr>
<tr>
<td>Error</td>
<td>12208.815</td>
<td>301</td>
<td>55.244</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>33477.343</td>
<td>305</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>24106.075</td>
<td>304</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .494 (Adjusted R Squared = .487)

Table 4.10  MCS-12 Difference Score Tests of Between Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>9459.552</td>
<td>3</td>
<td>3153.184</td>
<td>52.432</td>
<td>0.000</td>
<td>.416</td>
</tr>
<tr>
<td>Intercept</td>
<td>8116.353</td>
<td>1</td>
<td>8116.353</td>
<td>134.961</td>
<td>0.000</td>
<td>.379</td>
</tr>
<tr>
<td>MCS-12(baseline)</td>
<td>6205.392</td>
<td>1</td>
<td>6205.392</td>
<td>103.185</td>
<td>0.000</td>
<td>.318</td>
</tr>
<tr>
<td>Global Rating</td>
<td>3481.010</td>
<td>2</td>
<td>1740.505</td>
<td>28.942</td>
<td>0.000</td>
<td>.208</td>
</tr>
<tr>
<td>Error</td>
<td>13290.599</td>
<td>301</td>
<td>60.138</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>26010.986</td>
<td>305</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>22750.151</td>
<td>304</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .494 (Adjusted R Squared = .487)

RESPONSE RATE

Tables 4.11 and 4.12 present results comparing data for patients enrolled in the study to patients not enrolled for each of the practice guidelines. Patients were eligible for enrollment if an order was written for one of the antimicrobial therapies listed by either guideline between the dates of January 3, 2000 and March 26, 2000 and were able to complete the baseline health status survey.
From Table 4.11 thirty-six percent of patients ordered vancomycin during the study period were enrolled in the study. The reason most patients were not enrolled was due to inability to contact the patient while they were in the hospital. It should be noted that the vancomycin cohort was, in general, severely ill. This is shown by the large number of patients contacted unable to participate and refusal to participate. Additionally, the two groups were significantly different with respect to mortality (Enrolled = 8.2% versus Not Enrolled = 19.2%). Approximately twenty percent of the patients not enrolled in the study that received vancomycin therapy died during the hospital admission. Thus, severity of illness of the vancomycin group of patients in general was high. Of the patients enrolled, seven were lost to follow-up. Four patients were lost due to inability to contact them and three died post-discharge. The data for these seven patients was not used in the analyses that follow. Data for admission source, total cost and length of stay were not significantly different between the two groups.

From Table 4.12, twenty-two percent of the patients with drugs covered by the iv to po guideline were enrolled in the study. Once again, the main reason for not being enrolled in the study was due to not being contacted. A high number of patients were ordered therapies covered by the iv to po guideline, thus it was not possible to enroll them all due to resource limitations. There was significant difference between the groups with respect to admission source. More patients in the enrolled group were transferred to OSUMC (34.5% for enrolled and 23.7% for not enrolled). More patients in the not enrolled group were admitted by physician referral (19.3% for enrolled and
29.5% not enrolled). These result were not surprising. Transferred patients to OSUMC tend to be the highest acuity patients. Since patients enrolled tended to be in the hospital long enough to be contacted for enrollment, transferred patients may have been enrolled at a higher rate. However, mortality rate, total cost and total length of stay were not significantly different between enrolled and not enrolled groups.

Based on the results, it is difficult to determine if patients enrolled were different from patients not enrolled. One of the main outcome variables, cost, showed no significant difference between patients enrolled and not enrolled. Thus, for the purpose of this study, there did not appear to be significant selection bias for the outcome variables of interest.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Enrolled</th>
<th>Not Enrolled</th>
<th>Test Statistic</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>141</td>
<td>227</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Response Rate</td>
<td>38.3</td>
<td>61.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Reasons for Non-Enrollment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical/Mental Inability</td>
<td></td>
<td>21</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Patient Refusal</td>
<td></td>
<td>13</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Patient Not Contacted</td>
<td></td>
<td>148</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Reason Not Determined</td>
<td></td>
<td>42</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Admission Source – number (%)</td>
<td></td>
<td></td>
<td>0.707</td>
<td>3</td>
<td>NS</td>
</tr>
<tr>
<td>Clinic Referral</td>
<td>5 (3.7)</td>
<td>9 (3.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency Department</td>
<td>37 (27.6)</td>
<td>53 (22.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physician Referral</td>
<td>42 (31.4)</td>
<td>86 (36.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer</td>
<td>50 (37.3)</td>
<td>86 (36.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality Rate</td>
<td>11 (8.2)</td>
<td>45 (19.2)</td>
<td>13.67</td>
<td>133</td>
<td>0.034</td>
</tr>
<tr>
<td>Mean Total Cost - $ (+/- SD)</td>
<td>22,737 (1372)</td>
<td>24,937 (1734)</td>
<td>0.564</td>
<td>133</td>
<td>NS</td>
</tr>
<tr>
<td>Mean Total Length of Stay (+/- SD)</td>
<td>11.01 (8.40)</td>
<td>12.39 (10.97)</td>
<td>1.569</td>
<td>133</td>
<td>NS</td>
</tr>
<tr>
<td>Lost to Follow-Up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient Address and Phone Incorrect</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Post Discharge Death</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

a – the 7 patients lost to follow are not included in the rest of the data presented in the table or in future analyses
b – independent samples t-tests were used for continuous variables and chi-square tests of independence was used for categorical variables.
c – statistical significance is defined as a p < 0.05.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Enrolled(d)</th>
<th>Not Enrolled</th>
<th>Test Statistic(b)</th>
<th>df</th>
<th>Sig.(c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>180</td>
<td>645</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Response Rate</td>
<td>21.8</td>
<td>78.2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Reasons for Non-Enrollment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical/Mental Inability</td>
<td>-</td>
<td>14</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Patient Refusal</td>
<td>-</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Patient Not Contacted</td>
<td>-</td>
<td>428</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Reason Not Determined</td>
<td>-</td>
<td>209</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Admission Source – number (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinic Referral</td>
<td>3 (1.8)</td>
<td>36 (5.6)</td>
<td>15.97</td>
<td>3</td>
<td>0.003</td>
</tr>
<tr>
<td>Emergency Department</td>
<td>76 (44.4)</td>
<td>275 (42.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physician Referral</td>
<td>33 (19.3)</td>
<td>190 (29.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer</td>
<td>59 (34.5)</td>
<td>152 (23.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality Rate</td>
<td>9.4</td>
<td>9.3</td>
<td>11.53</td>
<td>170</td>
<td>NS</td>
</tr>
<tr>
<td>Mean Total Cost - $ (+/- SD)</td>
<td>17,893 (17,455)</td>
<td>16,804 (32,527)</td>
<td>0.422</td>
<td>170</td>
<td>NS</td>
</tr>
<tr>
<td>Mean Total Length of Stay – Days (+/- SD)</td>
<td>10.30 (8.21)</td>
<td>9.69 (14.19)</td>
<td>0.542</td>
<td>170</td>
<td>NS</td>
</tr>
<tr>
<td>Lost to Follow-Up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient Address and Phone Incorrect</td>
<td>8</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Post Discharge Death</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
</tr>
</tbody>
</table>

\(a\) – the 9 patients lost to follow are not included in the rest of the data presented in the table or in future analyses.

\(b\) – independent samples t-tests were used for continuous variables and chi-square tests of independence was used for categorical variables.

\(c\) – statistical significance is defined as \(p < 0.05\).
ORGANIZATION OF THE REMAINDER OF RESULTS

Up to this point, the results reported data pooled across the two respective guidelines. The remainder of the results split the two guideline results into separate analyses sections. In the first section, all of the results for the vancomycin guideline are reported. In the second section, all of the results for the iv to po guideline are reported. In each of the two sections, the data were presented as a comparison of the two study groups (e.g. CGP and NCGP). CGP was defined as patients treated with drug therapy orders compliant with guideline recommendations. Hence, NCGP was defined as patients treated with orders not compliant.

VANCOMYCIN GUIDELINE

The results for the vancomycin guideline are organized as follows: demographic and clinical history results are presented followed by the results of the four research questions.

Demographic Characteristics

Table 4.13 presents basic demographic data comparing CGP to NCGP. Three variables were found to be significantly different between groups including: age, admit source and admit service. The CGP were older patients on average than the NCGP (61.0 years old versus 55.1 years old). The CGP were admitted more through physician referral while the NCGP tended to be admitted more through the Emergency Department. Finally, CGP were admitted to a surgery service while the NCGP were
admitted to a medical service. The later result should be interpreted carefully since one of the approved uses for vancomycin is thoracic surgery prophylaxis. Of the 51 surgery service patients in the CGP, 49 were on the thoracic surgery service. There were no significant differences between the two groups for gender, marital status, education level, payer or race.
<table>
<thead>
<tr>
<th>Variable</th>
<th>CGP (n=82)</th>
<th>NCGP (n=52)</th>
<th>Test Statistic a</th>
<th>df</th>
<th>Sig. b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- std.dev.)</td>
<td>60.95 (14.48)</td>
<td>55.13 (15.26)</td>
<td>2.220</td>
<td>132</td>
<td>0.028</td>
</tr>
<tr>
<td>Gender #(% )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>59 (72.0)</td>
<td>29 (55.8)</td>
<td>3.696</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>Female</td>
<td>23 (28.0)</td>
<td>23 (44.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital Status #(% )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>49 (61.3)</td>
<td>25 (48.1)</td>
<td>2.220</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>Single</td>
<td>31 (38.8)</td>
<td>27 (51.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education Level #(% )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some H.S.</td>
<td>16 (19.5)</td>
<td>8 (15.7)</td>
<td>4.745</td>
<td>3</td>
<td>NS</td>
</tr>
<tr>
<td>H.S. Graduate</td>
<td>48 (58.5)</td>
<td>25 (49.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some College</td>
<td>12 (14.6)</td>
<td>8 (15.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College Graduate</td>
<td>6 (7.3)</td>
<td>10 (19.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payer #(% )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicaid</td>
<td>6 (7.3)</td>
<td>7 (13.5)</td>
<td>3.279</td>
<td>4</td>
<td>NS</td>
</tr>
<tr>
<td>Medicare</td>
<td>43 (52.4)</td>
<td>20 (38.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private – MCO</td>
<td>19 (20.2)</td>
<td>14 (26.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private – Other</td>
<td>8 (6.1)</td>
<td>5 (9.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>9 (11.0)</td>
<td>6 (11.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Admit Source #(% )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physician Referral</td>
<td>35 (42.7)</td>
<td>12 (23.1)</td>
<td>6.154</td>
<td>2</td>
<td>0.030</td>
</tr>
<tr>
<td>Emergency Dept.</td>
<td>18 (22.0)</td>
<td>19 (36.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer</td>
<td>29 (35.4)</td>
<td>21 (40.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Admit Service #(% )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td>51 (63.0)</td>
<td>13 (25.0)</td>
<td>18.282</td>
<td>1</td>
<td>0.005</td>
</tr>
<tr>
<td>Medicine</td>
<td>30 (37.0)</td>
<td>39 (75.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race #(% )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>72 (87.8)</td>
<td>39 (75.0)</td>
<td>3.670</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>Non-White</td>
<td>10 (12.2)</td>
<td>13 (25.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a - independent samples t-tests were used for continuous variables and chi-square tests of independence were used for categorical variables.
b - statistical significance is defined as a p < 0.05.

Table 4.13 Demographic Comparison of CGP and NCGP – Vancomycin
Clinical History Characteristics

Clinical history variables are proxy variables used to examine the acuity or 'severity' of patient illness in the two cohort groups. From Table 4.14 none of the clinical history variables were significantly different between the two groups. Variables examined included number of comorbid conditions, antibiotic use prior to admission, days of present illness, prior hospital admissions, prior emergency department visits, prior physician office visits, and baseline PCS-12 and MCS-12 scores.

<table>
<thead>
<tr>
<th>Variable</th>
<th>CGP(n=82)</th>
<th>NCGP(n=52)</th>
<th>Test Statistic&lt;sup&gt;a&lt;/sup&gt;</th>
<th>df</th>
<th>Sig&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comorbid Conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- Std.Dev.)</td>
<td>2.27 (1.32)</td>
<td>2.54 (1.60)</td>
<td>-1.060</td>
<td>132</td>
<td>NS</td>
</tr>
<tr>
<td>Prior Antibiotic Usage (%)*</td>
<td></td>
<td></td>
<td>2.053</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>Yes</td>
<td>11 (13.4)</td>
<td>11 (21.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>71 (86.6)</td>
<td>41 (78.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days of Present Illness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- Std.Dev.)</td>
<td>8.50 (10.98)</td>
<td>11.48 (11.63)</td>
<td>-1.497</td>
<td>132</td>
<td>NS</td>
</tr>
<tr>
<td>Prior Hosp Admit (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>38 (46.3)</td>
<td>23 (44.2)</td>
<td></td>
<td>0.666</td>
<td>1</td>
</tr>
<tr>
<td>No</td>
<td>44 (53.7)</td>
<td>29 (55.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency Dept. Visits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- Std.Dev.)</td>
<td>0.67 (0.87)</td>
<td>0.65 (0.59)</td>
<td>0.124</td>
<td>132</td>
<td>NS</td>
</tr>
<tr>
<td>Physician Office Visits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- Std.Dev.)</td>
<td>0.68 (1.36)</td>
<td>0.73 (1.10)</td>
<td>-0.206</td>
<td>130</td>
<td>NS</td>
</tr>
<tr>
<td>Baseline PCS-12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- Std.Dev.)</td>
<td>33.09 (9.74)</td>
<td>32.50 (8.99)</td>
<td>0.347</td>
<td>132</td>
<td>NS</td>
</tr>
<tr>
<td>Baseline MCS-12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- Std.Dev.)</td>
<td>41.36 (12.99)</td>
<td>39.11 (14.77)</td>
<td>0.927</td>
<td>132</td>
<td>NS</td>
</tr>
</tbody>
</table>

<sup>a</sup> - independent samples t-tests were used for continuous variables and chi-square tests of independence was used for categorical variables.

<sup>b</sup> - statistical significance is defined as a p < 0.05.

Table 4.14 Clinical History Characteristics – Vancomycin

190
Economic Outcome (Research Question One)

The economic variables examined in this study were: total hospital cost, drug therapy cost, laboratory cost, other cost, total length of stay (LOS), days of lost activity, hospital readmission, emergency department visits post discharge, physician office visits post discharge and new antibiotic therapy started post discharge (Table 4.15). All cost data is estimated cost using cost to charge ratios. Resource utilization data refers to the time period from the date of discharge to 30 days post discharge. No economic outcome variable was significantly different between the CGP and NCGP with exception of other costs (CGP=$14,705 and NCGP=$11,001). Other costs include all costs except room and board, drug therapy and laboratory cost. Hence, it included costs such as operating room costs, diagnostic test costs, physical therapy costs, etc. Thus, although other costs were significantly different between the groups, total costs were not different. The relative importance of other costs was difficult to determine.

Although there were no statistical differences between the two groups on all economic variables (with the exception of ‘other costs’ category), it should be noted that NCGP compared to CGP was higher by $2,062 on average. In addition, drug therapy costs were $891 higher on average and laboratory costs were $257 higher when comparing NCGP to CGP. Total LOS was higher in the NCGP (12.08 days) than the CGP (10.33). Thus, there are trends readily apparent from Table 4.15 in which the NCGP had higher costs.
Due to the large variation in cost, the data was log transformed to reduce the variation and the same analysis was performed. The transformed data remained non-significant and is not reported in the document.

<table>
<thead>
<tr>
<th>Variable</th>
<th>CGP (n=82)</th>
<th>NCGP (n=52)</th>
<th>Test Statistic*</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Hospital Cost</td>
<td>23,537 (15,397)</td>
<td>21,475 (16,691)</td>
<td>0.731</td>
<td>132</td>
<td>NS</td>
</tr>
<tr>
<td>Drug Therapy Cost</td>
<td>3,634 (4,158)</td>
<td>4,525 (5,533)</td>
<td>-0.997</td>
<td>132</td>
<td>NS</td>
</tr>
<tr>
<td>Laboratory Cost</td>
<td>2,077 (2,083)</td>
<td>2,334 (2,311)</td>
<td>-0.666</td>
<td>132</td>
<td>NS</td>
</tr>
<tr>
<td>Other Cost</td>
<td>14,705 (8,551)</td>
<td>11,001 (8,795)</td>
<td>2.417</td>
<td>132</td>
<td>0.017</td>
</tr>
<tr>
<td>Total LOS</td>
<td>10.33 (7.74)</td>
<td>12.08 (9.32)</td>
<td>-1.128</td>
<td>132</td>
<td>NS</td>
</tr>
<tr>
<td>Days of Lost Activity</td>
<td>4.40 (7.75)</td>
<td>6.30 (9.04)</td>
<td>-1.262</td>
<td>125</td>
<td>NS</td>
</tr>
<tr>
<td>Hospital Readmit #(%)</td>
<td></td>
<td></td>
<td>1.914</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>Yes</td>
<td>3 (3.9)</td>
<td>5 (10.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>74 (96.1)</td>
<td>45 (90.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency Dept. Visit</td>
<td>0.039 (0.19)</td>
<td>0.10 (0.42)</td>
<td>0.124</td>
<td>132</td>
<td>NS</td>
</tr>
<tr>
<td>Physician Office Visit</td>
<td>0.57 (1.01)</td>
<td>0.46 (0.73)</td>
<td>-0.206</td>
<td>130</td>
<td>NS</td>
</tr>
<tr>
<td>New Antibiotic #(%)</td>
<td></td>
<td></td>
<td>0.000</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>Yes</td>
<td>17 (22.1)</td>
<td>11 (22.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>60 (77.9)</td>
<td>39 (78.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a - independent samples t-tests were used for continuous variables and chi-square tests of independence was used for categorical variables.

b - statistical significance is defined as a p < 0.05.

Table 4.15 Economic Outcome - Vancomycin
Clinical Outcome (Research Question Two)

Table 4.16 presents clinical results observed for the vancomycin guideline. Both clinical variables studied were significantly different between the two groups. The CGP had a lower in-hospital mortality rate than the NCGP (CGP=2.4% and NCGP=17.3%). In addition, CGP had significantly higher success (73.2%) than the NCGP (48.1%). The latter effect was important since this was the effectiveness variable used in this study.

<table>
<thead>
<tr>
<th>Variable</th>
<th>CGP(n=82)</th>
<th>NCGP(n=52)</th>
<th>Test Statistic</th>
<th>df</th>
<th>Sig.</th>
<th>$\phi$</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality #(%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2 (2.4)</td>
<td>9 (17.3)</td>
<td>9.336$^d$</td>
<td>1</td>
<td>0.003</td>
<td>-.264</td>
<td>0.002</td>
</tr>
<tr>
<td>No</td>
<td>80 (97.6)</td>
<td>52 (82.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success #(%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>60 (73.2)</td>
<td>25 (48.1)</td>
<td>8.639$^d$</td>
<td>1</td>
<td>0.003</td>
<td>0.254</td>
<td>0.003</td>
</tr>
<tr>
<td>No</td>
<td>22 (26.8)</td>
<td>27 (51.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a - independent samples t-tests were used for continuous variables and chi-square tests of independence was used for categorical variables.

b - statistical significance is defined as $p < 0.05$.
c - Phi is a test of association
d - Due to small cell size, the data reported for mortality is the Fisher's Exact test.

Table 4.16 Clinical Outcome - Vancomycin

Health Status Outcome (Research Question Three)

Tables 4.17 to 4.19 indicate that order compliance did not have an effect on either follow-up health status scores or on difference score improvement between the two measurement periods. Tables 4.20 and 4.21 also showed order compliance had no effect on PCS-12 or MCS-12 difference scores after statistically controlling for baseline scores.
<table>
<thead>
<tr>
<th>Variable</th>
<th>CGP</th>
<th>NCGP</th>
<th>Test Statistic&lt;sup&gt;a&lt;/sup&gt;</th>
<th>df</th>
<th>Sig.&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Follow-Up PCS-12</td>
<td>39.14 (10.40)</td>
<td>39.92 (11.21)</td>
<td>-0.396</td>
<td>125</td>
<td>NS</td>
</tr>
<tr>
<td>Follow-Up MCS-12</td>
<td>44.10 (12.62)</td>
<td>44.33 (12.31)</td>
<td>-0.099</td>
<td>125</td>
<td>NS</td>
</tr>
</tbody>
</table>

<sup>a</sup> - independent samples t-tests.
<sup>b</sup> - statistical significance is defined as a p < 0.05.

Table 4.17 Follow-Up Health Status Outcome – Vancomycin

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGP</td>
<td>50</td>
<td>7.20</td>
</tr>
<tr>
<td>NCGP</td>
<td>77</td>
<td>6.70</td>
</tr>
<tr>
<td>Total</td>
<td>127</td>
<td>6.90</td>
</tr>
</tbody>
</table>

Table 4.18 Vancomycin PCS-12 Difference Score Descriptives

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGP</td>
<td>50</td>
<td>5.12</td>
</tr>
<tr>
<td>NCGP</td>
<td>77</td>
<td>3.32</td>
</tr>
<tr>
<td>Total</td>
<td>127</td>
<td>4.03</td>
</tr>
</tbody>
</table>

Table 4.19 Vancomycin MCS-12 Difference Score Descriptives

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum f Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>2199.54</td>
<td>2</td>
<td>1099.77</td>
<td>12.07</td>
<td>.000</td>
<td>.163</td>
</tr>
<tr>
<td>Intercept</td>
<td>4358.76</td>
<td>1</td>
<td>4358.76</td>
<td>47.85</td>
<td>.000</td>
<td>.278</td>
</tr>
<tr>
<td>PCS-12(baseline)</td>
<td>2191.97</td>
<td>1</td>
<td>2191.97</td>
<td>24.07</td>
<td>.000</td>
<td>.163</td>
</tr>
<tr>
<td>Order Compliance</td>
<td>11.80</td>
<td>1</td>
<td>11.80</td>
<td>0.13</td>
<td>.720</td>
<td>.001</td>
</tr>
<tr>
<td>Error</td>
<td>11294.67</td>
<td>124</td>
<td>91.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>19538.39</td>
<td>127</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>13494.22</td>
<td>126</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.20 PCS-12 Difference Score Tests of Between Subjects Effects (ANCOVA) – Vancomycin
<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>3686.04</td>
<td>2</td>
<td>1843.02</td>
<td>21.36</td>
<td>.000</td>
<td>.256</td>
</tr>
<tr>
<td>Intercept</td>
<td>5166.13</td>
<td>1</td>
<td>5166.13</td>
<td>59.87</td>
<td>.000</td>
<td>.326</td>
</tr>
<tr>
<td>MCS-12 (baseline)</td>
<td>3587.38</td>
<td>1</td>
<td>3587.38</td>
<td>41.58</td>
<td>.000</td>
<td>.251</td>
</tr>
<tr>
<td>Order Compliance</td>
<td>42.76</td>
<td>1</td>
<td>42.76</td>
<td>0.50</td>
<td>.483</td>
<td>.004</td>
</tr>
<tr>
<td>Error</td>
<td>10699.46</td>
<td>124</td>
<td>86.29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>16448.04</td>
<td>127</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>14385.50</td>
<td>126</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.21  MCS-12 Difference Score Tests of Between Subjects Effects (ANCOVA) - Vancomycin

**Cost Effectiveness (Research Question Four)**

Table 4.22 showed the average CE ratios and the incremental CE ratio. It should be noted, the calculation of total costs for the CGP equals the cost of producing the outcome for the patient plus the cost of implementing the guideline. The cost of implementation was estimated at $15,625 and included the time of the expert, printing and distribution costs. This estimate was based on the budget amount spent to produce 32 guidelines at the OSUMC. The assumption was made that equal resources were spent on the 32 guidelines approved.

The result indicated that for the vancomycin guideline, $8,980 is needed to produce one additional successful antimicrobial therapy (e.g. incremental analysis). The cost is higher and the effect (benefit) is higher.
<table>
<thead>
<tr>
<th>Alternatives</th>
<th>Costs(C)</th>
<th>Effectiveness (E) Per 100 patients</th>
<th>C/E Ratio ($ per successful antimicrobial therapy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCPG</td>
<td>2,147.500</td>
<td>48</td>
<td>44,740</td>
</tr>
<tr>
<td>CPG</td>
<td>2,372.000</td>
<td>73</td>
<td>32,493</td>
</tr>
<tr>
<td>Increment (of CPG over NCGP)</td>
<td>224,500</td>
<td>25</td>
<td>8,980</td>
</tr>
</tbody>
</table>

Table 4.22  Cost Effectiveness – Vancomycin Guideline

**Vancomycin Guideline Conclusion**

The results of the vancomycin guideline order compliance analyses indicate that increased effect was observed. No differences between the groups were found for economic or health status outcomes.

**IV TO PO GUIDELINE**

The results for the iv to po guideline are organized as follows: demographic and clinical history is presented followed by the results of the four research questions.

**Demographic Characteristics**

Table 4.23 showed the results of demographic characteristics when comparing the CGP cohort to the NCGP cohort. One demographic variable was significantly different between the two groups – admit service. More CGP (83.0%) were admitted to medical services than to surgical services (NCGP=17.0%).
<table>
<thead>
<tr>
<th>Variable</th>
<th>CGP(n=107)</th>
<th>NCGP(n=64)</th>
<th>Test Statistic</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- std.dev.)</td>
<td>56.79 (17.59)</td>
<td>55.36 (18.25)</td>
<td>0.506</td>
<td>169</td>
<td>NS</td>
</tr>
<tr>
<td>Gender #(%),</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>62 (57.9)</td>
<td>35 (54.7)</td>
<td>0.173</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>Female</td>
<td>45 (42.1)</td>
<td>29 (45.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital Status #(%),</td>
<td></td>
<td></td>
<td>0.002</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>Married</td>
<td>43 (41.0)</td>
<td>26 (40.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>62 (59.0)</td>
<td>38 (59.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education Level #(%),</td>
<td></td>
<td></td>
<td>3.235</td>
<td>3</td>
<td>NS</td>
</tr>
<tr>
<td>Some H.S.</td>
<td>28 (26.7)</td>
<td>15 (23.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H.S. Graduate</td>
<td>51 (48.6)</td>
<td>25 (39.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some College</td>
<td>13 (12.4)</td>
<td>13 (20.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College Graduate</td>
<td>13 (12.4)</td>
<td>11 (17.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payer#(%),</td>
<td></td>
<td></td>
<td>0.770</td>
<td>3</td>
<td>NS</td>
</tr>
<tr>
<td>Medicaid</td>
<td>15 (14.0)</td>
<td>8 (12.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicare</td>
<td>54 (50.5)</td>
<td>29 (45.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private</td>
<td>22 (20.6)</td>
<td>16 (25.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>16 (15.0)</td>
<td>11 (17.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Admit Source #(%),</td>
<td></td>
<td></td>
<td>6.406</td>
<td>2</td>
<td>NS</td>
</tr>
<tr>
<td>Physician Referral</td>
<td>16 (15.0)</td>
<td>20 (31.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency Dept.</td>
<td>51 (47.7)</td>
<td>25 (39.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer</td>
<td>40 (37.4)</td>
<td>19 (29.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Admit Service #(%),</td>
<td></td>
<td></td>
<td>9.032</td>
<td>1</td>
<td>0.007</td>
</tr>
<tr>
<td>Surgery</td>
<td>18 (17.0)</td>
<td>24 (37.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td>88 (83.0)</td>
<td>40 (62.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race #(%),</td>
<td></td>
<td></td>
<td>1.191</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>White</td>
<td>70 (65.4)</td>
<td>47 (73.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-White</td>
<td>37 (34.6)</td>
<td>17 (26.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a - independent samples t-tests were used for continuous variables and chi-square tests of independence was used for categorical variables.

b - statistical significance is defined as p < 0.05.

Table 4.23  Demographic Comparison of CGP and NCGP – IV to PO
Clinical History Characteristics

The only significant difference between the two groups for clinical history variables was Emergency Department utilization prior to being admitted to the hospital (Table 4.24). The CGP cohort had more mean visits (1.00) than the NCGP (0.67) prior to being admitted to the hospital.

<table>
<thead>
<tr>
<th>Variable</th>
<th>CGP(n=107)</th>
<th>NCGP(n=64)</th>
<th>Test Statistic&lt;sup&gt;a&lt;/sup&gt;</th>
<th>df</th>
<th>Sig.&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comorbid Conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- Std.Dev.)</td>
<td>2.39 (1.41)</td>
<td>2.33 (1.26)</td>
<td>0.300</td>
<td>169</td>
<td>NS</td>
</tr>
<tr>
<td>Prior Antibiotic Usage #(%):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>30 (28.0)</td>
<td>12 (18.8)</td>
<td>1.864</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>No</td>
<td>77 (72.0)</td>
<td>52 (81.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days of Present Illness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- Std.Dev.)</td>
<td>13.43 (11.11)</td>
<td>10.43 (11.33)</td>
<td>1.709</td>
<td>169</td>
<td>NS</td>
</tr>
<tr>
<td>Prior Hosp Admit #(%):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>54 (50.5)</td>
<td>29 (45.3)</td>
<td>0.426</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>No</td>
<td>53 (49.5)</td>
<td>35 (54.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency Dept. Visits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- Std.Dev.)</td>
<td>1.00 (0.84)</td>
<td>0.67 (0.87)</td>
<td>2.443</td>
<td>169</td>
<td>0.016</td>
</tr>
<tr>
<td>Physician Office Visits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- Std.Dev.)</td>
<td>1.01 (1.58)</td>
<td>0.87 (1.79)</td>
<td>0.517</td>
<td>167</td>
<td>NS</td>
</tr>
<tr>
<td>Baseline PCS-12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- Std.Dev.)</td>
<td>31.55 (7.81)</td>
<td>33.74 (8.74)</td>
<td>-1.689</td>
<td>168</td>
<td>NS</td>
</tr>
<tr>
<td>Baseline MCS-12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (+/- Std.Dev.)</td>
<td>38.10 (14.06)</td>
<td>38.47 (14.57)</td>
<td>-0.161</td>
<td>168</td>
<td>NS</td>
</tr>
</tbody>
</table>

<sup>a - independent samples t-tests were used for continuous variables and chi-square tests of independence was used for categorical variables.</sup>

<sup>b - statistical significance is defined as a p < 0.05.</sup>
Economic Outcome (Research Question One)

For the iv to po guideline (Table 4.25), none of the economic outcome variables examined were different between the CGP and NCGP. Hence, complying with the guideline or not complying appears to show no effect on any economic outcome. One would expect drug cost to be lower in the CGP assuming enough switches from iv to po occurred. Although not statistically significant, total costs for the NCGP are $1,250 higher than the CGP on average. Also, total LOS is shorter for the NCGP as compared to the CGP. Thus, the average daily cost for the CGP is $1,559 while for the NCGP it is $2,113.

As for the vancomycin guideline, the cost data was log transformed due to the large variation noted in the results. Again, the results remained non-significant.
<table>
<thead>
<tr>
<th>Variable</th>
<th>CGP(n=107)</th>
<th>NCGP(n=64)</th>
<th>Test Statistic(^a)</th>
<th>df</th>
<th>Sig.(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Hospital Cost</td>
<td>17,425 (16,119)</td>
<td>18,675 (19,596)</td>
<td>-0.452</td>
<td>169</td>
<td>NS</td>
</tr>
<tr>
<td>Drug Therapy Cost</td>
<td>3,608 (4,780)</td>
<td>3,532 (4,556)</td>
<td>0.103</td>
<td>168</td>
<td>NS</td>
</tr>
<tr>
<td>Laboratory Cost</td>
<td>2,074 (2,181)</td>
<td>1,833 (2,053)</td>
<td>0.714</td>
<td>168</td>
<td>NS</td>
</tr>
<tr>
<td>Other Cost</td>
<td>8,719 (8,795)</td>
<td>11,358 (11,666)</td>
<td>-1.562</td>
<td>168</td>
<td>NS</td>
</tr>
<tr>
<td>Total LOS</td>
<td>11.18 (8.31)</td>
<td>8.84 (7.88)</td>
<td>1.811</td>
<td>169</td>
<td>NS</td>
</tr>
<tr>
<td>Days of Lost Activity</td>
<td>5.85 (9.39)</td>
<td>6.14 (8.10)</td>
<td>-0.197</td>
<td>154</td>
<td>NS</td>
</tr>
<tr>
<td>Hospital Readmit #(%)</td>
<td>9 (9.2)</td>
<td>2 (3.4)(^c)</td>
<td>1.829</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>89 (90.8)</td>
<td>56 (96.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency Dept. Visit</td>
<td>0.10 (0.42)</td>
<td>0.017 (0.13)</td>
<td>1.858</td>
<td>154</td>
<td>NS</td>
</tr>
<tr>
<td>Physician Office Visit</td>
<td>0.62 (1.12)</td>
<td>0.36 (0.64)</td>
<td>1.850</td>
<td>154</td>
<td>NS</td>
</tr>
<tr>
<td>New Antibiotic #(%)</td>
<td>24</td>
<td>12</td>
<td>0.296</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>74</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) - independent samples t-tests were used for continuous variables and chi-square tests of independence were used for categorical variables.

\(^b\) - statistical significance is defined as a p < 0.05.

\(^c\) - Due cells with low counts, Fisher's Exact Test is reported.

Table 4.25  
Economic Outcome – IV to PO

200
Clinical Outcome (Research Question Two)

Table 4.26 showed the results of the clinical outcome of order compliance.

There were no significant clinical outcome differences between the CGP and the NCGP. Mortality shows a trend which favors the CGP over the NCGP (94.4% survival versus 84.4% survival) but remain statistically non significant.

<table>
<thead>
<tr>
<th>Variable</th>
<th>CGP (n=107)</th>
<th>NCGP (n=64)</th>
<th>Test Statistic</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital Mortality</td>
<td></td>
<td></td>
<td>4.738</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>Yes</td>
<td>6 (5.6)c</td>
<td>10 (15.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>101 (94.4)</td>
<td>54 (84.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success</td>
<td></td>
<td></td>
<td>0.032</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>Yes</td>
<td>65 (60.7)</td>
<td>38 (59.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>42 (39.3)</td>
<td>26 (40.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a - independent samples t-tests were used for continuous variables and chi-square tests of independence was used for categorical variables.
b - statistical significance is defined as a p < 0.05.
c - Due to small cell size, the data reported for mortality is the Fisher’s Exact test.

table 4.26 Clinical Outcome – IV to PO

Health Status Outcome (Research Question Three)

Significant differences were not found between the CGP and NCGP on either of the two subscales for the follow-up health status survey (see Table 4.27). Tables 4.28 and 4.29 present no significant difference on the change scores for either the PCS-12 or MCS-12, respectively. Even after the effect of baseline score was controlled for using ANCOVA (Tables 4.30 and 4.31), no significant effect was found for order compliant group. Again, the result may have been influenced by the psychometric properties of the SF-12. However, ‘floor’ effect in the guideline was not a significant concern.
<table>
<thead>
<tr>
<th>Variable</th>
<th>CGP Mean</th>
<th>NCGP Mean</th>
<th>Test Statistic</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Follow-Up PCS-12</td>
<td>38.11(11.31)</td>
<td>41.06(9.48)</td>
<td>-1.670</td>
<td>154</td>
<td>NS</td>
</tr>
<tr>
<td>Follow-Up MCS-12</td>
<td>42.23(12.52)</td>
<td>44.16(12.24)</td>
<td>-0.938</td>
<td>154</td>
<td>NS</td>
</tr>
</tbody>
</table>

a - independent samples t-tests.
b - statistical significance is defined as a p < 0.05.

Table 4.27 Follow-Up Health Status Outcome – IV to PO

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGP</td>
<td>5.66</td>
<td>9.84</td>
</tr>
<tr>
<td>NCGP</td>
<td>6.92</td>
<td>10.26</td>
</tr>
<tr>
<td>Total</td>
<td>6.53</td>
<td>10.12</td>
</tr>
</tbody>
</table>

Table 4.28 IV to PO PCS-12 Difference Score Descriptives

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGP</td>
<td>3.32</td>
<td>9.82</td>
</tr>
<tr>
<td>NCGP</td>
<td>5.23</td>
<td>10.63</td>
</tr>
<tr>
<td>Total</td>
<td>4.63</td>
<td>10.39</td>
</tr>
</tbody>
</table>

Table 4.29 IV to PO MCS-12 Difference Score Descriptives

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum f Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>1706.14</td>
<td>2</td>
<td>853.07</td>
<td>9.22</td>
<td>.000</td>
<td>.108</td>
</tr>
<tr>
<td>Intercept</td>
<td>3305.38</td>
<td>1</td>
<td>3305.38</td>
<td>35.71</td>
<td>.000</td>
<td>.189</td>
</tr>
<tr>
<td>PCS-12(baseline)</td>
<td>1652.53</td>
<td>1</td>
<td>1652.53</td>
<td>17.86</td>
<td>.000</td>
<td>.105</td>
</tr>
<tr>
<td>Order Compliance</td>
<td>2.35</td>
<td>1</td>
<td>2.35</td>
<td>0.025</td>
<td>.874</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>14160.84</td>
<td>153</td>
<td>92.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>22509.25</td>
<td>156</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>15866.98</td>
<td>155</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.30 PCS-12 Difference Score Tests of Between Subjects Effects (ANCOVA) – IV to PO
From Table 4.32, the incremental cost of the CGP over the NCGP is $12,000. The incremental effect is two. Thus, the incremental cost effectiveness ratio was calculated to be $6,000. Again, the domination rule was not met. The cost of the CGP was higher than that for the NCGP. No claim of cost effectiveness is possible for this result in this study.

### Cost-Effectiveness Analysis (Research Question Four)

From Table 4.32, the incremental cost of the CGP over the NCGP is $12,000. The incremental effect is two. Thus, the incremental cost effectiveness ratio was calculated to be $6,000. Again, the domination rule was not met. The cost of the CGP was higher than that for the NCGP. No claim of cost effectiveness is possible for this result in this study.
IV to PO Guideline Conclusion

Although there were few statistically significant effects on outcome found between the two groups for the iv to po guideline, this was not a completely negative result. Compliance with the iv to po guideline recommendation does appear to produce better economic results than non-compliance.
REFERENCES FOR CHAPTER 4


8. Unpublished Data. The Ohio State University Medical Center. Columbus, OH; The Ohio State University; 1997.

205
CHAPTER 5

DISCUSSION, IMPLICATIONS, AND RECOMMENDATIONS

This Chapter begins with a brief review of the basis for the study. Then, the results of the study are discussed and interpreted in relation to the study rationale. This part of the discussion was organized around the guideline evaluation exercise and the four research questions posed in the study. In each section, in addition to discussion, implication of the results are presented. The Chapter ends with limitations to the research and suggestions for future research based on current study results.

BASIS OF STUDY

The basis for this study is allied with the paucity of literature on the effect of clinical practice guidelines on patient outcomes.\(^1\,^2\) Lack of literature evidence was especially true for guidelines developed locally.\(^3\,^4\) At OSUMC there had not been any study of implemented practice guidelines. The vast majority of published evaluation studies focus on length of stay or process type variables.\(^5\,^6\) Few studies were found
which examined cost and outcome data. The studies that contained this information were conducted on national guidelines or length of stay guidelines.\textsuperscript{5-7} Thus, local development and lack of research on the effects of clinical practice guidelines on patient outcomes form the basis for study need.

The second basic underpinning for this study was the issue of following guideline recommendations. If developing practice guidelines meet defined needs for an institution, then evaluation results should show positive effect. Hence, following recommendations suggested in a practice guideline should be studied relative to not following guideline recommendations. The issue of compliance has been studied in the literature.\textsuperscript{8,10} Therefore, compliance with guideline recommendations was utilized in the study as a method to define alternatives.

Finally, the basis of this study is on justifying the use of resources. The development and implementation of practice guidelines consume significant resources.\textsuperscript{11,12} Prudence and accountability necessitates the need to determine if the resources are used for benefit.

GUIDELINE EVALUATION

The guideline evaluation exercise conducted in this study found the OSUMC practice guideline development process robust. Though only 60.0\% and 64.0\% of the twenty-five development standards were judged by practitioners as met for the iv to po and vancomycin guidelines, respectively, the results are in line with other published evaluations of guideline development.\textsuperscript{13} Shaneyfelt et al. found that for published, peer-
reviewed guidelines, overall mean adherence was 43.1%. For standards on development and format, the two OSUMC guidelines each met 8 of 10 standards or 80.0%. In the Shaneyfelt et al. study, published guidelines met 51.1% of these standards. For the evidence identification and summary standards the two OSUMC guidelines were judged to have met 5 of 10 while in the Shaneyfelt et al. study the published guidelines met 33.6% of these standards. Finally, for the formulation of recommendations standards, the OSUMC iv to po met 2 of 5 while the vancomycin met 3 of 5. In the Shaneyfelt et al. study published guidelines met 46.0 percent.

<table>
<thead>
<tr>
<th>Standards</th>
<th>OSUMC (%)</th>
<th>Shaneyfelt et al. (%)</th>
<th>Test Statistic</th>
<th>Df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development and format (standards 1 to 10)</td>
<td>80.0/80.0</td>
<td>51.1</td>
<td>16.345</td>
<td></td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Evidence identification and summary (standards 11 to 20)</td>
<td>50.0/50.0</td>
<td>33.6</td>
<td>8.005</td>
<td></td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Formulation of recommendations (standards 21 to 25)</td>
<td>60.0/40.0</td>
<td>46.0</td>
<td>4.261/0.7826</td>
<td></td>
<td>&lt;0.05/NS</td>
</tr>
<tr>
<td>All Standards</td>
<td>64.0/60.0</td>
<td>43.1</td>
<td>10.135/6.627</td>
<td></td>
<td>&lt;0.05</td>
</tr>
</tbody>
</table>

a - χ² test of goodness-of-fit
b - Critical value of χ² = 3.841 at α = 0.05.
c - NS

Table 5.1  Development Standards Met – OSUMC versus Published Results

The analysis in Table 5.1 indicate that the two OSUMC guidelines were better than those in the published literature. Thus, the issue of guideline quality is not of
concern for this study. For future studies of practice guidelines, it is highly
recommended that this issue be examined. The issue of guideline quality is especially
important for locally developed guidelines that are produced using various untested
methodologies. It is unlikely that development methods will be standardized. Thus, the
next best way to examine development quality is to examine the end product with
respect to the list of development standard presented by Shaneyfelt et al.\textsuperscript{13}

PSYCHOMETRIC PROPERTIES

The psychometric properties of the SF-12 for use in acute and severely ill
patients suffering from infection provided encouraging results. Although relationships
between the PCS-12 and MCS-12 were in the expected direction and were significant,
additional testing of this property in the study population is recommended. Validity is
never proven, rather, information is consistently gathered over time that supports the
notion of valid measurement. In general, this study found evidence to indicate that the
two SF-12 summary subscales were valid for use in this population.

The PCS-12 and MCS-12 appear to be reliable measures of health status in the
population. Internal consistency (e.g., Cronbach's $\alpha$) was calculated for the PCS-12
and MCS-12 subscales for baseline and follow-up measurements, respectively. The
Cronbach $\alpha$ for the PCS-12 was 0.6411 at baseline and 0.6888 at follow-up. The
Cronbach $\alpha$ for the MCS-12 at baseline was 0.7053 and at follow-up 0.6844. For group
level health status measurement, Nunnally et al. recommends a reliability coefficient of
0.70.\textsuperscript{14} Thus, the internal consistency was slightly below this standard reported in
literature. A concern regarding reliability in the study population was the fact that significant floor effects might occur and were expected for baseline PCS-12 and MCS-12. However, the ‘floor’ effect in the follow-up was within the guideline of 15% as defined by McHorney et al. Thus, the ‘floor effect’ was not a significant concern for the study.

The responsiveness of both subscales of the SF-12 performed well in two responsiveness measures studied. Since a major concern for all generic instruments is the ability to detect change, this is a significant result found in the study. It has been argued in the literature that disease specific instruments should be used to improve responsiveness. However, the PCS-12 and MCS-12 were responsive to change in this study population. The standardized response mean was high in this study (PCS-12 at baseline 0.75). In addition, the instrument captured change at a rate higher than the conditions studied by the developers. For example, Stewart et al. reports a 7.6 point difference in PCS-36 is equivalent to the effect of having diabetes, a clinically meaningful difference. Ware et al. reports that mean difference scores for patients with heart valve replacement surgery (before and after) was 7.64 for the PCS-36. For the current study mean difference score for the PCS-12 was 6.45 (see Table 4.7). Similarly, for the MCS-36, Ware et al notes a mean change score of 3.86 for patients diagnosed and treated for clinical depression for one year. For the current study, mean differences for the MCS-12 was 3.81 (see Table 4.8).

The results from the global rating of health change by respondents support the SF-12 responsiveness. Even when patients report their health as remaining the same,
the SF-12 measured improvement (mean change = 5.88 for the PCS-12 and 1.59 for the MCS-12). When patients reported improvement, the SF-12 captured this change also (mean change scores 13.18 for the PCS-12 and 8.55 for the MCS-12).

The SF-12 met the psychometric requirements for use in group level monitoring. The instrument performed as it was expected and captured change between measurements. However, no significant differences were found between the CGP and the NCGP groups in terms of health outcomes measured by the SF-12. This results can be interpreted to mean that that the decision to write an order following a guideline or not following did not make a difference in health status between the two groups as measured by the SF-12.

RESPONSE RATE

There did not appear to be significant selection bias issues in the study. However, one major limitation of the study design is the inability to ensure the study groups were similar at baseline. The concern revolves around the issue of enrolling 'sicker' patients in one of the groups (CGP versus NCGP) and not the other. More severely ill patients were hospitalized for longer periods of time than the less 'ill' patients. A consequence related to the study design of being in the hospital longer is that the probability of being enrolled might have been increased. This concern is not a serious threat since being in the hospital longer should increase the probability of being enrolled for either group – CGP and NCGP.
From the available data, selection bias did not appear to be significant issue other than the high rate of mortality in the vancomycin not enrolled group. Almost twenty percent of this group died while in the hospital. Selection bias due to high mortality could not be checked in this study.

VANCOMYCIN RESULTS

First reporting demographic and clinical history characteristics followed by the research questions and concluding implications for all of the research questions present the results for the guidelines.

Demographic Characteristics

The significant differences found between CGP and NCGP on age and admission source are discussed in detail since these two variables might explain variance in outcome variables discussed later in the section. The older CGP as compared to the NCGP may indicate more vulnerable patients are treated with orders compliant with the vancomycin guideline. In other words, patients who have defined need for vancomycin usually receive it. Additional support for the argument that CGP were more vulnerable is based upon the higher proportion admitted to OSUMC (e.g., admission service) via a referring physician. Since the physician prior to admission may have known these patients, a clear need for vancomycin may have been established. Transferred patients and emergency department patients tend to be more complicated (increased clinical uncertainty) patients and to be in the NCGP (almost 77% of the NCGP fall into one of
these two categories). Uncertainty about the clinical background of a patient may lead
to higher utilization of vancomycin even when it is not indicated. These two variables
appear to corroborate the possibility that complying with vancomycin guideline
recommendations was higher for well known and vulnerable patients.

The third significant variable, admit service, was not unexpected. One approved
(or compliant) use of vancomycin was for prophylaxis for thoracic surgery patients.
Thus, all thoracic surgery patients who received vancomycin were in the CGP. This
result is not deemed clinically important for interpretation of study results.

Clinical History Characteristics

While demographic variables discussed above indicate more vulnerable patients
may be treated following recommendations, no clinical history variable was
significantly different between CGP and NCGP. Thus, there is some evidence that the
two groups did not differ clinically at baseline. In other words, drug therapy orders for
guideline covered therapies do not appear to be written based on recent clinical history.
This was the intention of the practice guideline. The result was interpreted in two ways.
First, the interpretation was made that both groups of patients are severely ‘ill’. The
second was that there was no difference in illness between the groups. Both
explanations support the contention that CGP were ordered vancomycin following
guideline recommendations for vulnerable conditions based on demographic
characteristics rather than current clinical condition. A similar argument could be made for the NCGP. For example, due to a patient's severity of illness, vancomycin may be improperly ordered.

**Economic Outcome (Research Question One)**

Of the economic variables, 'other costs', was significantly different for CGP and NCGP. There were too many components of 'other' costs to determine why this difference exists. Looking only at the economic variables, it might be tempting to conclude that following or not following the guideline recommendations does not matter from an economic perspective. However, caution must be used when making this interpretation since no clinical effect has been examined. While it is true that there is no difference between CGP and NCGP in terms of cost, benefit produced by CGP versus NCGP has not been examined. Finding no effect for the economic variables between the two groups might be interpreted to mean that the more vulnerable CGP group produced positive economic results. The economic results were no worse for the CGP versus the NCGP despite the fact there was evidence the CGP were more vulnerable at enrollment.

**Clinical Outcome (Research Question Two)**

Perhaps the most important result found in the vancomycin study involved the clinical outcomes. Mortality was significantly lower in the CGP (2.4%) versus NCGP (17.3%). As discussed earlier, this result may be due to selection. In addition, the
antimicrobial therapy success rate was higher for CGP group (73.2% versus 48.1%).

Despite the lack of observed differences for economic outcomes, there were significant clinical outcomes associated with CGP. Also, the CGP were older and treated at similar cost, yet, better clinical outcomes were obtained. The results discussed thus far support the need for efforts to educate to move more patients from the NCGP to the CGP. For similar cost, additional benefits may accrue by getting more patients in the CGP. The clinical results are not independent results. In fact, mortality is one condition defined as antimicrobial failure.

**Health Status Outcomes (Research Question Three)**

There was no effect of order compliance on the raw score of either subscales of the SF-12. The severe nature of the patient’s illnesses in both groups may have limited the ability of the SF-12 to capture results that differentiate the groups. However, as discussed earlier, the SF-12 measured and captured change in this population. The issue is that the rate of change between the groups was not different. One cause of the result might be that, on average, both groups improved. In fact, the CGP improved 7.20 points between the two measurements for the PCS-12 and the NCGP improved 6.70 points. Similar results were found for the MCS-12 (5.12 improvement for the CGP and 3.32 for NCGP). An alternative explanation is that although evidence was found to support the responsiveness of the SF-12 in the study population, it may not have detected subtle differences between the groups. Perhaps a disease specific instrument
for infectious disease processes would capture differential rates of change between the
groups.\textsuperscript{16} A third plausible explanation may be that the measures of compliance were
not well specified which contaminated the groups.

**Cost-Effectiveness (Research Question Four)**

The cost-effectiveness results were very limited in interpretation. As discussed
in Chapter 2, only the domination rule would be used to label one alternative as cost-
effective. There was no global budget or set price to use as possible rules. Therefore,
although the $8,980 needed to produce one additional antimicrobial therapy success in
the CGP cohort might seem low and reasonable, the results did not dominate the NCGP
cohort. To dominate, the cost would need to be lower. For the vancomycin guideline,
costs were higher and effect was higher. Thus, the decision makers at OSUMC must
answer the question of whether the benefit is worth the cost. It should be noted that
Kaplan et al. suggests a program or alternative with a cost effectiveness ratio less than
$20,000 should be implemented.\textsuperscript{30} No methodology was employed in the study to
answer this question. Costs were higher due to the added estimated development and
implementation costs for the CGP.

**Vancomycin Conclusion**

The results of the research questions for the vancomycin guideline indicate there
were benefits accrued by the CGP. The CGP cohort was older, yet had better clinical
outcomes at a similar cost to the NCGP. However, the result did not prove to be ‘cost-
effective' by definition. The interpretation of the results should not be inferred to mean the CGP cohort did not show significant benefits. It is clear there were significant benefits. Unfortunately, there is no way to determine if the benefits are ‘worth’ the cost at OSUMC for the purposes of this study.

IV TO PO GUIDELINE

The results of the iv to po guideline are discussed and interpreted in the following order: demographic characteristics, clinical history characteristics, economic outcome, clinical outcome, health status outcome and cost-effectiveness analysis.

Demographic Characteristics

Only one demographic variable was significantly different between the CGP and NCGP. More CGP were admitted to medical services as opposed to surgical services. Although this was a significant result, it may reflect the fact that medical services in general use more antimicrobial therapy or were more aware of the practice guideline. In fact, the Infectious Disease Service is a medical service and authored both of the guidelines in the study. Thus, awareness and adoption issues may have influenced the difference.

Clinical History Characteristics

The only significant clinical history characteristic, which differed between the CGP and NCGP, was Emergency Department utilization rate prior to hospital
admission. The CGP cohort utilized more Emergency Department visits than the NCGP prior to being admitted. This result indicated, that, as in the vancomycin guideline analysis, the CGP cohort might have been more acutely 'ill' than the NCGP. Since this cohort used Emergency Department visits more frequently this interpretation is plausible. However, no other clinical history variable corroborates the result. Thus, this interpretation was made with caution.

Economic Outcome (Research Question One)

None of the economic variables examined were statistically significant for difference between the CGP and NCGP. As in the discussion for vancomycin, this result tempts the interpretation that following or not following guideline recommendation had no effect on economic outcome. Unlike the vancomycin discussion there were no explanations based on demographic or clinical history variables. What makes the no effect explanation even more enticing was the fact that the sole purpose of the iv to po guideline was to effect drug therapy costs. The results of this study note no effect on the total drug therapy cost between CGP and NCGP.

Even though no statistical significance was found, for the economic variables, the results hint that there was an operationally significant component. From table 4.25, it was noted that total costs for the NCGP were $1,250 higher on average than for the CGP. This economic factor must be examined carefully. From the non-response analysis (Table 4.12) 645 patients per quarter (or 2,580 patients annually) receive an order for one of the medications covered by the recommendations of the iv to po
 guideline. Assuming the same rate of non-compliance occurs (approximately 40%), then 1,032 patient would be in the NCGP annually. By multiplying 1,032 patients by $1,250 annually, a cost of $1,290,000 in cost may be incurred. Thus, there is a highly operationally significant issue associated with iv to po compliance.

By examining the guideline closely, it should be noted that the CGP consist of two types of patients. First, patients who should not be switched to po medication are included in the CGP. In addition, patients started on iv and switched to po and who are clinically supposed to be switched are included in CGP. Similarly, in the NCGP there are two types of patients. First, there were patients who should have been switched but were not. In addition, there were patients who were switched too early and should not have been switched and fail (patient is ordered new iv therapy).

Furthermore, the results from Table 4.25 indicate that a negative relationship exists between total costs and total length of stay (LOS). For example, total cost is $1,250 higher for the NCGP. At the same time, total LOS is lower (8.84 days for NCGP versus 11.18 days for CPG. Thus higher cost in fewer days may not be optimal.
Figure 5.1 is a useful to study the logic behind decisions being made. The diagram has been adopted from other applications. If a patient was switched and should have been switched then that is appropriate and is what is supposed to happen. A similar appropriate decision is not switching a patient who should not be switched. The problems begin to occur when patients are switched and they should not have been due to severity issues (Type I error). The Type I scenario may lead to higher cost. In addition, the patient may deteriorate clinically leading to a cycle of negative economic and clinical events. This is the scenario physicians are trained to avoid. Physicians do not want their patient to be in a Type I situation. Thus, physicians over-compensate by putting patients who should be switched on iv therapy and leaving them on iv therapy until they are successful (Type II error). This is the explanation of how the NCGP has

<table>
<thead>
<tr>
<th>Should Patient Have Been Switched</th>
<th>Patients Was Actually Switched</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Appropriate Switch</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Type I Error</td>
</tr>
</tbody>
</table>

| Appropriately Switched Patients | Type II Error |

| Appropriately Non-Switched Patients | Type I Error |

Figure 5.1 Guideline Compliance Decision Making
higher costs with lower lengths of stay. Physicians have real reasons for purposefully committing Type II error. They ensure the patient will get better quicker, they reduce professional liability risk and generally limit their accountability. The problem arises when Type II error occurs in large numbers, costs may quickly increase. Type II error is the most common reason found for non-compliance in this study (less than 10 cases of failure occurred due to switching too soon).

Thus, Type II error explains the paradox or balance between Type I and Type II error that must occur when deciding to follow or not follow guideline recommendations. Incentives may need to be employed or disincentives may need to be removed to move physicians from making Type II error to appropriate decisions.

Clinical Outcome (Research Question Two)

Clinical effect was not different between CPG and NCPG for the iv to po guideline. Thus, the clinical outcomes were not influenced by the decision to follow the guideline or not follow the guideline. The implication may be that the outcomes of using oral medication versus intravenous medication therapies does not show any negative clinical effect in this study. This was a positive result. Intravenous therapy is often used due to concern for the severity of the infectious process. The result of this study may refute this notion, given appropriate guidance is given as to when to switch a patient safely from iv to po.
Health Status Outcome (Research Question Three)

Order compliance was found to have no effect on health status follow-up or difference scores as measured by the SF-12. Again, for the iv to po guideline, this is a positive result. Switching appropriately had no detrimental effect on health status by definition.

Cost Effectiveness Analysis (Research Question Four)

Cost effectiveness was not found for the iv to po guideline by definition in Chapter 2. Cost was higher for the CGP as compared to the NCGP due to the addition of guideline development and implementation costs. Average cost was not significantly different between the two groups. The result was interesting given the fact that cost reduction was the goal of the guideline. However, operational significance has been discussed and is important. The no difference in effect was the desired outcome for the iv to po guideline. However, the lack of lower cost eliminated the possibility of using the domination rule. Thus, the conclusion for the iv to po guideline was that there is no way to state that compliance with guideline recommendations were cost effective by definition. However, operationally significant economic cost difference is extremely important to the bottom line of the institution. Thus, it is up to the decision-makers at OSUMC to determine effective strategies to increase compliance, especially when Type II error is occurring.
LIMITATIONS

There were four major limitations that have been identified throughout the research. First, a recurrent theme throughout the study is the issue of selection bias. The study design used did not allow randomization of patients to groups or treatment. Thus, there was no sure method to ensure the groups were equal. Selected demographic and clinical history variables did not indicate significant differences existed between enrolled and not enrolled patients. Mortality rate was higher in the none enrolled patient population. The result is a consequence of patients having to be able physically to complete the baseline survey. Additionally, there is no evidence that suggests that mortality would have impacted one study group over the other (CGP versus NCGP).

Second, the results of the study are limited in generalizability. The results are limited to locally developed guidelines. There is need to study local guidelines. Yet, this narrow focus limits the ability to infer outside of a single institution. Additionally, the results may be limited in terms of generalizability to drug therapy guidelines. The results may not be useful for the study of other types of guidelines that may not deal with drug therapy. The results may not even be generalizable to other drug therapy classes. Despite the limitations in generalizability noted, the methodology used in the study is generalizable to other studies conducted on a variety of guidelines in many different settings.

Third, due to lost to follow-up and a higher than expected compliance rate, power is low. Many of the results presented show trends, but remain not significant. This may be due to the low number of patients in the NCGP. A larger sample would have increased the power of the analysis.

223
Finally, the results of the study rely on the ability of the research to classify a patient as having been treated following or not following guideline recommendations. The measures of compliance may not have been defined well enough. A good example of this issue is a timing effect. The guidelines are not specific of when a patient should be switched from iv to po therapy. The vancomycin guideline also does not specify length of required therapy. Thus, the classification of CGP and NCGP may be ambiguous. The result may be confounding which may have limited the effects found in the study.

FUTURE RESEARCH

Potential for research on practice guidelines in general and on outcome evaluation is unlimited. As the number of guidelines produced expand, more research is needed to determine the effect. The fact that many guidelines are being developed and implemented at the local level highlight a need for research. The results of this study indicate there are some positive effects associated with following guideline recommendations.

More research is needed on which examine guideline quality. Some standardization is necessary. Without standardization in quality, it may difficult to compare the results of guideline evaluation studies. This study notes significant limits in generalizability. Currently, most studies on guideline evaluation note the same limitations as the current study with respect to generalizability.
Replication studies within the same institution are needed especially as new implementation procedures are utilized. One example currently being implemented is the use of computerized reminders in the physician order entry system. These reminders may move more patients from the Type II error situation discussed earlier to an appropriate decision for the iv to po guideline. The impact of these new implementation techniques must be studied.

In addition, replication studies on the use of the SF-12 health status instrument need to be conducted on the acute, infectious disease population. The results of the current study support the general psychometric properties of the instrument in this population. Additional studies would be useful for validation purposes.

Finally, more studies need to be conducted on the iv to po guideline to examine the Type II error discussed. The clinical and physician specific reasons for not switching patients to oral antimicrobial therapy need to be conducted. Research is needed to determine the costs of patients who have Type I error discussed. If a patient fails the switch, then the additional costs need to be examined.
REFERENCES FOR CHAPTER 5


APPENDIX A

Converting Intravenous to Oral Antimicrobials
The Ohio State University Medical Center
Clinical Practice Guideline

Converting Intravenous to Oral Antimicrobials

Definitions of Recommendation Levels
(The quality of evidence determines the strength of the guideline.)

<table>
<thead>
<tr>
<th>Level of Evidence</th>
<th>Quality of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Good</td>
</tr>
<tr>
<td>II</td>
<td>Fair</td>
</tr>
<tr>
<td>III</td>
<td>Poor</td>
</tr>
</tbody>
</table>

**Level of Evidence**
- **I**: Good Evidence obtained from at least one properly randomized controlled trial or from well-designed cohort or case-control analytic studies, preferably from more than one center or research group.
- **II**: Fair Evidence obtained from multiple time series with or without intervention, or dramatic results in uncontrolled experiments.
- **III**: Poor Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees.

**Strength of Recommendation**

<table>
<thead>
<tr>
<th>Strength</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>It is advised that this be done... in a periodic health examination, or a diagnostic maneuver should be done, or treatment of a medical program should be given. There is good evidence (level I) to support the recommendation.</td>
</tr>
<tr>
<td>B</td>
<td>This may be done... in a previous health examination, or a diagnostic maneuver may be done, or treatment of a medical problem may be given.</td>
</tr>
<tr>
<td>B1</td>
<td>It should be done in most cases. There is fair evidence (level II) to support the recommendation.</td>
</tr>
<tr>
<td>B2</td>
<td>It should NOT be done in most cases. There is poor evidence (level III) regarding the recommendation, which may be made on other grounds.</td>
</tr>
<tr>
<td>C</td>
<td>It is advised that this NOT be done... in a periodic health examination, or a diagnostic maneuver should not be done, or treatment of a medical problem should not be given. There is good evidence (level I) to support the recommendation that this not be done.</td>
</tr>
</tbody>
</table>

Note: Practice guidelines are not standards that are meant to be applied rigidly and followed in virtually all cases. Patient choice and physician judgment must remain central to the selection of diagnostic tests and therapy. Practice guidelines should be helpful to physicians and patients alike in making their decisions.

A Leadership Council for Clinical Value Enhancement Initiative

229
Objective

To maintain optimal treatment, reduce patient discomfort, reduce adverse reactions from intravenous lines, and reduce costs.

Guideline Statements

("A" Recommendation: It is advised that this be done. There is good evidence [level I] to support the recommendation.)

1. For some drug-organism combinations, oral administration is so reliable that it can be used as soon as gastrointestinal function is satisfactory and the patient is stable. For others, it is best to wait until recovery is underway.

2. For some drug-organism combinations, there is no oral equivalent, and converting from intravenous to oral may be impossible or require a change in drug class.

3. With some infections, such as endocarditis or meningitis, switching is usually not appropriate, and intravenous therapy is needed for the full course of treatment.

4. When converting from intravenous to oral therapy at the time of hospital discharge, it is not necessary or even useful to delay discharge and observe a patient on oral medication for 24 hours to detect relapse. Tolerance to the oral drug can be determined by giving a dose on the evening prior to discharge. Note: As indicated, patients may be discharged on intravenous antimicrobials.

Key Words

antimicrobials, endocarditis, gastrointestinal function, infections, intravenous, meningitis, oral therapy.

RECOMMENDATIONS

A Leadership Council for Clinical Value Enhancement Initiative
### Equivalent with Normal GI Function

*"A" recommendation: It is advised that this be done.
There is good evidence (level I) to support the recommendation.*

<table>
<thead>
<tr>
<th>Drug</th>
<th>IV Dose</th>
<th>Cost</th>
<th>Oral Dose</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>azithromycin</td>
<td>500 mg q24h</td>
<td>$$</td>
<td>500 mg load, then 250 mg qd</td>
<td>$$$</td>
</tr>
<tr>
<td>ciprofloxacin</td>
<td>200 mg q12h</td>
<td>$$</td>
<td>250 mg bid</td>
<td>$$$</td>
</tr>
<tr>
<td></td>
<td>400 mg q12h</td>
<td>$$$</td>
<td>500 mg bid</td>
<td>$$$</td>
</tr>
<tr>
<td></td>
<td>400 mg q8h</td>
<td>$$$$</td>
<td>750 mg bid</td>
<td>$$$$</td>
</tr>
<tr>
<td>TMP/SMZ</td>
<td>2.5 mg TMP/kg q12h</td>
<td>$</td>
<td>1 DS (160/800 mg) tab bid</td>
<td>$</td>
</tr>
<tr>
<td></td>
<td>5 mg TMP/kg q8h</td>
<td>$</td>
<td>2 DS (160/800 mg) tabs tid</td>
<td>$</td>
</tr>
<tr>
<td>fluconazole</td>
<td>200 mg q24h</td>
<td>$$$$</td>
<td>200 mg qd</td>
<td>$$$$</td>
</tr>
<tr>
<td></td>
<td>400 mg q24h</td>
<td>$$$$$</td>
<td>400 mg qd</td>
<td>$$$$$</td>
</tr>
</tbody>
</table>

### Nearly Equivalent

*"B1" Recommendation: This should be done in most cases.
There is fair evidence (level II) to support the recommendation.*

<table>
<thead>
<tr>
<th>Drug</th>
<th>IV Dose</th>
<th>Cost</th>
<th>Oral Dose</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>clindamycin</td>
<td>300-900 mg q8h</td>
<td>$</td>
<td>150-300 mg qid</td>
<td>$$$</td>
</tr>
<tr>
<td>doxycycline</td>
<td>100 mg q12-24h</td>
<td>$</td>
<td>100 mg bid or qd</td>
<td>$</td>
</tr>
<tr>
<td>metronidazole</td>
<td>500 mg q6h</td>
<td>$</td>
<td>500 mg qid</td>
<td>$</td>
</tr>
</tbody>
</table>

### Adequate for Many Infections

(Serum Concentration Significantly Higher with Intravenous Administration)

*"B1" Recommendation: This should be done in most cases.
There is fair evidence (level II) to support the recommendation.*

<table>
<thead>
<tr>
<th>Drug</th>
<th>IV Dose</th>
<th>Cost</th>
<th>Oral Dose†</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>penicillin G</td>
<td>1 million units q6h</td>
<td>$</td>
<td>penicillin V 500 mg qid</td>
<td>$</td>
</tr>
<tr>
<td>ampicillin</td>
<td>1 g q6h</td>
<td>$</td>
<td>amoxicillin 500 mg tid</td>
<td>$</td>
</tr>
<tr>
<td>ampicillin/</td>
<td>1.5 g q6h</td>
<td>$$</td>
<td>amoxicillin/clavulanate 875 mg bid</td>
<td>$$$$</td>
</tr>
<tr>
<td>sulbactam</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nafcillin</td>
<td>1 g q6h</td>
<td>$</td>
<td>dicloxacillin 500 mg qid</td>
<td>$$$</td>
</tr>
<tr>
<td>cefazolin</td>
<td>1 g q8h</td>
<td>$</td>
<td>cephalexin 500 mg qid</td>
<td>$</td>
</tr>
<tr>
<td>cefepime or</td>
<td>1 g q24h</td>
<td>$$</td>
<td>cefpodoxime 200 mg bid or</td>
<td>$$$$</td>
</tr>
<tr>
<td>ceftriaxone</td>
<td>1 g q12h</td>
<td>$$</td>
<td>cefuroxime 250-500 mg bid</td>
<td>$$$$</td>
</tr>
<tr>
<td>erythromycin</td>
<td>500 mg q6h</td>
<td>$</td>
<td>erythromycin 500 mg qid</td>
<td>$</td>
</tr>
</tbody>
</table>

† For penicillin and cephalosporins, a similar, but not identical drug must be used.
Reference


A Leadership Council for Clinical Value Enhancement Initiative
Suggested Evaluation/Outcomes Measure

If resources were available, the following would be appropriate:

- Percent of patients actually switched as outlined.
- Percent cure, failure, relapse.
- Incidence phlebitis and line sepsis
- Direct drug costs.
- Nursing time.
APPENDIX B

Criteria for Use of Vancomycin
### Criteria for Use of Vancomycin (IV and Oral)

#### Definitions of Recommendation Levels

*(The quality of evidence determines the strength of the guideline.)*

<table>
<thead>
<tr>
<th>Quality of Evidence</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Good</td>
<td>Evidence obtained from at least one properly randomized controlled trial or from well-designed cohort or case-control analytic studies, preferably from more than one center or research group.</td>
</tr>
<tr>
<td>II Fair</td>
<td>Evidence obtained from multiple time series with or without intervention, or dramatic results in uncontrolled experiments.</td>
</tr>
<tr>
<td>II Poor</td>
<td>Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strength of Recommendation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A <strong>It is advised that this be done...</strong></td>
<td>in a periodic health examination, or a diagnostic maneuver should be done, or treatment of a medical problem should be given. There is good evidence (level I) to support the recommendation.</td>
</tr>
<tr>
<td>B <strong>This may be done...</strong></td>
<td>in a periodic health examination, or a diagnostic maneuver may be done, or treatment of a medical problem may be given.</td>
</tr>
<tr>
<td>B1 <strong>It should be done in most cases.</strong></td>
<td>There is fair evidence (level II) to support the recommendation.</td>
</tr>
<tr>
<td>B2 <strong>It should NOT be done in most cases.</strong></td>
<td>There is poor evidence (level III) regarding the recommendation, which may be made on other grounds.</td>
</tr>
<tr>
<td>C <strong>It is advised that this NOT be done...</strong></td>
<td>in a periodic health examination, or a diagnostic maneuver should not be done, or treatment of a medical problem should not be given. There is good evidence (level I) to support the recommendation that this not be done.</td>
</tr>
</tbody>
</table>

**Note:** Practice guidelines are not standards that are meant to be applied rigidly and followed in virtually all cases. Patient choice and physician judgment must remain central to the selection of diagnostic tests and therapy. Practice guidelines should be helpful to physicians and patients alike in making their decisions.

*A Leadership Council for Clinical Value Enhancement Initiative*
Objective

To promote the prudent use of vancomycin and help prevent infections due to vancomycin-resistant *Enterococcus faecium* (VRE).

Guideline Statement

(A recommendation: It is advised that this be done. There is good evidence [level I] to support the recommendation)

These recommendations give indications and contraindications for prophylaxis and treatment with intravenous vancomycin and for treatment and dosing with oral vancomycin.

Key Words

antimicrobials, *Clostridium difficile* diarrhea, endocarditis, *Enterococcus faecium* infection, metronidazole, prophylaxis, vancomycin.

Background

Vancomycin is a very important antibiotic. Intravenously, it is extensively used to prevent and treat infections caused by β-lactam-resistant gram-positive cocci (such as methicillin-resistant staphylococci). Orally, it is extensively used to treat antibiotic-associated *Clostridium difficile* diarrhea. Unfortunately, there has been an increase in infections caused by VRE. These infections represent a serious therapeutic challenge because they often occur in compromised patients, and there are no good therapeutic options. In addition, the genetic material responsible for vancomycin resistance can be transferred to staphylococci, and the potential threat of vancomycin-resistant staphylococci emerging in patients is a major concern.

*A Leadership Council for Clinical Value Enhancement Initiative*
INTRAVENOUS VANCOMYCIN

(These are "A" recommendations: It is advised that this be done. There is good evidence [level I] to support the recommendation.)

Prophylaxis With Intravenous Vancomycin

Indications

- Prophylaxis as recommended by the American Heart Association guidelines for patients undergoing specified diagnostic or surgical procedures and who are at high risk for endocarditis (1-2 doses).*

- Prophylaxis for surgical implantation of prosthetic materials such as artificial heart valves, vascular grafts and joint implants. A single 1 g dose should be administered just prior to surgery, and repeated if the procedure lasts more than 6 hours (2 doses maximum). If an infection is discovered at the time of surgery, a prescription for therapy should be written postoperatively.

* See also Clinical Practice Guidelines for Bacterial Endocarditis Prophylaxis.

Treatment With Intravenous Vancomycin

Indications

- Infections caused by or suspected to be caused by β-lactam-resistant gram-positive organisms (e.g., methicillin-resistant staphylococci and ampicillin-resistant enterococci). If used empirically, treatment should be changed if the infection subsequently proves to be caused by an organism or organisms susceptible to other appropriate antimicrobials.

- Serious infections caused by β-lactam-susceptible gram-positive organisms in patients with a serious β-lactam allergy (e.g., enterococcal endocarditis in a patient with penicillin allergy).
Intravenous Vancomycin Should NOT Be Used:

- For routine surgical prophylaxis. Cefazolin 1g (1-3 doses) or cefotetan 1g (1 dose only) is appropriate.
- For prophylaxis to prevent colonization or infection of intravascular catheters.
- For prophylaxis in patients on continuous ambulatory peritoneal dialysis.
- To eradicate methicillin-resistant *Staphylococcus aureus* (MRSA) colonization.
- To treat infections caused by β-lactam-susceptible gram-positive organisms (e.g., methicillin-susceptible staphylococci and ampicillin-susceptible enterococci) unless there is serious β-lactam allergy and no other appropriate alternative therapy. β-lactams are more effective.
- For prolonged empiric treatment when cultures are consistently negative for β-lactam-resistant gram-positive organisms.
- To treat patients with renal failure when the only justification is the convenience of infrequent dosing.
- For empiric treatment of febrile neutropenic, surgical, or any other patients unless the illness is immediately life-threatening and there is reasonable probability that the infection is caused by a β-lactam-resistant gram-positive organism. The presence of a central venous catheter alone is insufficient justification unless there is an exit-site or tunnel infection.
- To treat a patient with a single positive line culture for methicillin-resistant *Staphylococcus epidermidis* (MRSE), or any other organism that is part of the normal skin flora and of low virulence, when the patient is not acutely ill or when there have been multiple negative cultures drawn during the same period. Ideally, infection due to these organisms should be confirmed by a positive culture from a peripheral vein.
Dosing for Intravenous Vancomycin

The usual daily adult intravenous dose is 2 grams divided as 1g every 12 hours. Dose adjustments should be made with reduced creatinine clearance.

<table>
<thead>
<tr>
<th>Creatinine Clearance</th>
<th>Dose Adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-80 ml/min</td>
<td>q 24-72 hr</td>
</tr>
<tr>
<td>10-50 ml/min</td>
<td>q 72-240 hr</td>
</tr>
<tr>
<td>&lt; 10 ml/min</td>
<td>q 240 hr</td>
</tr>
</tbody>
</table>

ORAL VANCOMYCIN

(These are "A" recommendations: It is advised that this be done. There is good evidence [level I] to support the recommendation.)

Treatment With Oral Vancomycin

Indications

- Diarrhea due to *C. difficile* toxin ONLY if the disease is severe and potentially life-threatening, if there is a contraindication to prescribing oral metronidazole, or if the patient has failed a previous course of metronidazole (250 mg po qid).

Dosing for Oral Vancomycin

- The usual dose to treat diarrhea due to *C. difficile* toxin is 125 mg po qid for 10 days. Higher doses or longer courses of treatment should be used ONLY for treatment failure on the usual dose.
Algorithm

Not applicable.
Background Information/Supportive Literature

This guideline is adapted from the following:


Adapted by

Robert Fass, MD; Department of Internal Medicine; Division of Infectious Diseases

Copyright

© 1996, The Ohio State University. No part of this publication may be reproduced in any form without permission in writing from The Ohio State University.

Approved by

Suggested Evaluation/Outcomes Measures

- Frequency of infections caused by *Enterococcus faecium*.
- Frequency of vancomycin-resistance among strains of *Enterococcus faecium*.
- Frequency of vancomycin-resistance among strains of *Staphylococcus aureus*, *Staphylococcus epidermidis* and *Staphylococcus haemolyticus*.
APPENDIX C

List of Study Variables
<table>
<thead>
<tr>
<th>Variable Type</th>
<th>Variable Name</th>
<th>Operational Definition</th>
<th>Measure ment Level</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient Demographic</td>
<td>Age</td>
<td>years at study enrollment</td>
<td>Interval</td>
<td>Medical Record</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>Male; Female</td>
<td>Nominal</td>
<td>Medical Record</td>
</tr>
<tr>
<td></td>
<td>Admission Service</td>
<td>Medical Service; Surgical Service; Other</td>
<td>Nominal</td>
<td>Medical Record</td>
</tr>
<tr>
<td></td>
<td>Primary Payer Type</td>
<td>Medicaid; Medicare; Private -Managed Care; Private -Other; Self-Pay; No Pay</td>
<td>Nominal</td>
<td>Medical Record</td>
</tr>
<tr>
<td></td>
<td>Marital Status</td>
<td>Married; Single</td>
<td>Nominal</td>
<td>Medical Record</td>
</tr>
<tr>
<td></td>
<td>Employment Status (on admission)</td>
<td>Employed Full Time; Employed Part Time; Not Employed; Student; Retired</td>
<td>Nominal</td>
<td>First Patient Survey</td>
</tr>
<tr>
<td></td>
<td>Education Level</td>
<td>Some High School; High School Graduate; Some College; College Graduate; Advanced College Graduate</td>
<td>Ordinal</td>
<td>First Patient Survey</td>
</tr>
<tr>
<td>Clinical History</td>
<td>Primary Diagnosis (DRG)</td>
<td>DRG</td>
<td>Nominal</td>
<td>Financial Database</td>
</tr>
<tr>
<td></td>
<td>ICD-9 Codes</td>
<td>List of 10 codes by order of clinical importance</td>
<td>Nominal</td>
<td>Financial Database</td>
</tr>
<tr>
<td></td>
<td>Comorbid Conditions</td>
<td>Presence of the following: Diabetes; renal disease (acute/chronic); AIDS; cancer (active/history); organ transplant; cardiovascular disease (MI/CHF); pulmonary disease (asthma/COPD)</td>
<td>Nominal</td>
<td>Medical Record</td>
</tr>
<tr>
<td></td>
<td>Admission Source</td>
<td>Emergent; Scheduled; Transfer</td>
<td>Nominal</td>
<td>Medical Record</td>
</tr>
<tr>
<td></td>
<td>Days of Present Illness</td>
<td>Number of Days in previous 30 not able to perform normal daily activities due to infection</td>
<td>Interval</td>
<td>First Patient Survey</td>
</tr>
<tr>
<td>Process Variables</td>
<td>At admission or in hospital prior to study enrollment, did the patient receive immunosuppressive drug therapy:</td>
<td>Nominal</td>
<td>Medical Record</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>---------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>Antimicrobial Drug Usage</td>
<td>At admission or in hospital prior to study enrollment, did the patient receive antimicrobial drug therapy:</td>
<td>Nominal</td>
<td>Medical Record</td>
<td></td>
</tr>
<tr>
<td>Previous Hospital Admission</td>
<td>Number of admissions in 30 days prior to enrollment for infection treatment.</td>
<td>Interval</td>
<td>First Patient Survey</td>
<td></td>
</tr>
<tr>
<td>ED Utilization</td>
<td>Number of ED in 30 days prior to enrollment for infection treatment.</td>
<td>Interval</td>
<td>First Patient Survey</td>
<td></td>
</tr>
<tr>
<td>Physician Office Visit Utilization</td>
<td>Number of physician office visits in 30 days prior to enrollment for infection treatment.</td>
<td>Interval</td>
<td>First Patient Survey</td>
<td></td>
</tr>
<tr>
<td>Fever</td>
<td>Highest temperature reading in first 24 hours after study enrollment.</td>
<td>Interval</td>
<td>Medical Record</td>
<td></td>
</tr>
<tr>
<td>White Blood Cell (WBC) count</td>
<td>Highest WBC count in first 24 hours after study enrollment.</td>
<td>Interval</td>
<td>Medical Record</td>
<td></td>
</tr>
<tr>
<td>Pulse Rate</td>
<td>Highest pulse rate in first 24 hours after study enrollment.</td>
<td>Interval</td>
<td>Medical Record</td>
<td></td>
</tr>
<tr>
<td>Baseline Health Status</td>
<td>SF-12 item scores and summary scale scores.</td>
<td>Interval</td>
<td>First Patient Survey</td>
<td></td>
</tr>
<tr>
<td>Process Variables</td>
<td>Endocarditis: Yes No</td>
<td>Nominal</td>
<td>Medical Record</td>
<td></td>
</tr>
<tr>
<td>Process Variables</td>
<td>Meningitis: Yes No</td>
<td>Nominal</td>
<td>Medical Record</td>
<td></td>
</tr>
<tr>
<td>Process Variables</td>
<td>GI Functioning (Diet): patient takes normal diet: Yes No</td>
<td>Nominal</td>
<td>Medical Record</td>
<td></td>
</tr>
<tr>
<td>Process Variables</td>
<td>GI Functioning (oral drug therapy): patient takes oral medicines: Yes No</td>
<td>Nominal</td>
<td>Medical Record</td>
<td></td>
</tr>
<tr>
<td>Process Variables</td>
<td>Purpose of Therapy: Prophylactic Treatment</td>
<td>Nominal</td>
<td>Medical Record</td>
<td></td>
</tr>
<tr>
<td>Process Variables</td>
<td>Thoracic Surgery: Yes No</td>
<td>Nominal</td>
<td>Medical Record</td>
<td></td>
</tr>
<tr>
<td>Process Variables</td>
<td>B-lactam resistant gram + organisms (empirical): Yes No</td>
<td>Nominal</td>
<td>Medical Record</td>
<td></td>
</tr>
<tr>
<td>Process Variables</td>
<td>B-lactam allergy: Yes No</td>
<td>Nominal</td>
<td>Medical Record</td>
<td></td>
</tr>
<tr>
<td>Economic Outcomes</td>
<td>B-lactam resistant gram + organism (lab confirmed)</td>
<td>Yes</td>
<td>Nominal</td>
<td>Medical Record</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------------------------------</td>
<td>-----</td>
<td>---------</td>
<td>----------------</td>
</tr>
<tr>
<td></td>
<td>Total Hospital Cost</td>
<td>dollars</td>
<td>Interval</td>
<td>Financial Database</td>
</tr>
<tr>
<td></td>
<td>Overall Length of Stay (LOS)</td>
<td>days</td>
<td>Interval</td>
<td>Financial Database</td>
</tr>
<tr>
<td></td>
<td>Post Enrollment Length of Stay (LOS)</td>
<td>days</td>
<td>Interval</td>
<td>Financial Database</td>
</tr>
<tr>
<td></td>
<td>Drug Therapy Cost</td>
<td>dollars</td>
<td>Interval</td>
<td>Financial Database</td>
</tr>
<tr>
<td></td>
<td>Laboratory Cost</td>
<td>dollars</td>
<td>Interval</td>
<td>Financial Database</td>
</tr>
<tr>
<td></td>
<td>Days of Lost Activity</td>
<td>Number of Days in last 30 after date of discharge not able to perform normal daily activities due to infection</td>
<td>Interval</td>
<td>Second Patient Survey</td>
</tr>
<tr>
<td></td>
<td>Hospital Readmissions</td>
<td>Number of readmissions for infection occurring 30 days starting the day of discharge.</td>
<td>Interval</td>
<td>Second Patient Survey</td>
</tr>
<tr>
<td></td>
<td>ED Visits</td>
<td>Number of ED visits for infection occurring 30 days starting the day of discharge.</td>
<td>Interval</td>
<td>Second Patient Survey</td>
</tr>
<tr>
<td></td>
<td>Physician Office Visits</td>
<td>Number of office visits for infection occurring 30 days starting the day of discharge.</td>
<td>Interval</td>
<td>Second Patient Survey</td>
</tr>
<tr>
<td></td>
<td>New Antimicrobial Therapy</td>
<td>Number of patients obtaining new antimicrobial therapy from a physician office visit occurring 30 days starting the day of discharge.</td>
<td>Interval</td>
<td>Second Patient Survey</td>
</tr>
<tr>
<td>Clinical Outcomes</td>
<td>Mortality</td>
<td>Number of In-Hospital Deaths</td>
<td>Interval</td>
<td>Medical Record</td>
</tr>
<tr>
<td></td>
<td>Drug Therapy Failures</td>
<td>Number of Failures</td>
<td>Interval</td>
<td>Medical Record</td>
</tr>
<tr>
<td>Health Status (SF-12)</td>
<td>SF-12</td>
<td>Physical Component Summary</td>
<td>Interval</td>
<td>First and Second Patient Surveys</td>
</tr>
<tr>
<td></td>
<td>SF-12</td>
<td>Mental Component Summary</td>
<td>Interval</td>
<td>First and Second Patient Surveys</td>
</tr>
<tr>
<td></td>
<td>SF-12</td>
<td>Item Analysis</td>
<td>Interval</td>
<td>First and Second Patient Surveys</td>
</tr>
</tbody>
</table>
APPENDIX D

Baseline Patient Cover Letter
John Doe  
1112 Street  
Columbus, Oh  

Dear,

We are currently conducting a short, mailed survey research project to measure how well patients recently treated at The Ohio State University Medical Center feel. This survey is part of a research project designed to examine the decisions made with respect to health care services provided and the impact of these decisions on our patients. It is very important that health care providers begin to examine how the decisions we make impact the patients we care for. For a more detailed description of the research, please read the enclosed informed consent form.

Since the survey is part of a research project, informed consent must be obtained. Please note that participation in the study is confidential and voluntary. If you decide to participate in the study, please sign the second page of the consent form and return it along with the completed survey. If you do not wish to participate in the study, you may simply throw the information away. All information obtained from the study is kept strictly confidential. No information shared by any respondent will be given to your physician or anyone else. All data will be reported as group (over all participants) information.

We have made a significant effort to keep the survey short (less than five minutes) to reduce the amount of your time involved in the study. Participating in the study requires the following:

1. Sign the informed consent form (where marked in red ink).
2. Complete the initial survey.
3. Return the initial survey in the postage paid envelope.
4. Complete a second, similar survey in two to four weeks (which will be mailed to you).

In 2-4 weeks a second, similar survey will be sent to you. It is extremely important that each participant complete the current or initial survey and the second, follow up survey in two to four weeks. To participate in the research it will take approximately 10 minutes of your time.

Your consideration to participate in this study is greatly appreciated! If you have any questions or concerns with participating in the study, do not hesitate to call me at 614 292-1716.

Sincere Thanks,

Dennis W. Grauer, M.S.  
Phone: 614-292-1716
APPENDIX E

Patient Consent Form
CONSENT TO INVESTIGATIONAL TREATMENT OR PROCEDURE

I, _____________________________, hereby authorize or direct Dr. Dev S. Pathak associates or assistants of his choosing, to perform the following treatment or procedure.

To conduct a brief (five to seven minutes) written survey upon myself.

The experimental portion of the treatment or procedure is:

The completion of the written survey.

This is done as part of an investigation entitled:

Cost, Outcomes and Physician Compliance with Clinical Practice Guidelines: A Case Study of Vancomycin Use and Intravenous to Oral Antimicrobial Guidelines at The Ohio State University Medical Center.

1. Purpose of the procedure or treatment:

To measure the general health of the patient. Current economic pressure on the health care system and the demand for quality has led to the development of clinical practice guidelines. Clinical practice guidelines are tools designed to help in specific health care decisions. The study purpose is to measure the effect of compliance with locally developed and implemented drug therapy guidelines. The study will examine the cost, clinical effect and patient effect (i.e., the general health of the patient) to determine the cost-effectiveness of compliance with clinical guidelines. Thus, the evaluation of general health over time is an important patient outcome variable for the study. Basic to the research is examining the impact of medical decisions on the health perception of the patient.

2. Possible appropriate alternative procedure or treatment (not to participate in the study is always an option):

None. The patient may choose to complete the survey or not complete the survey. In no case will the care received by the patient be affected by the decision to participate or not participate in the study.

3. Discomforts and risks reasonably to be expected:

No physical discomfort nor risk is involved in the study. No direct treatment or procedure is involved in participating in the study. The main discomfort for participating patients is completing the survey and the time necessary to complete this task (approximately 5 to 7 minutes).

No subject will be identified in publications or reports of the study and all information collected will be reported only in aggregate form. Data will be collected from the following three sources: the patient survey, the patient medical record and the hospital fiscal database. All data will be kept confidential. Procedures used to ensure confidentiality include the use of two independent databases. In the first database, patient specific data will be temporarily housed. This data will include the patient’s medical record number, name, mailing address and a random five digit
numerical code. This is the information required for collection of the 30 day follow up survey. The management of this data base and the collection of the 30 day follow up survey will be conducted by an individual not connected to the research project other than financial payment for services. After the collection of all patient research data, the database with patient identifiers will be destroyed. The second database will serve as the main research data bank and will include no patient specific information.

4. Possible benefits for subjects/society:

For the participating subject, the study does not have a direct benefit. The study objective is to evaluate the effect of clinical practice guidelines. Specifically, is compliance with practice guideline recommendations that focus on drug therapy cost-effective? Clinical practice guidelines are documents created for the purpose of making better decisions with respect to the health care options selected. The answer may be important to society because of the large amount of time and resources expended to develop practice guidelines.

5. Anticipated duration of subject's participation (including number of visits):

The subject's participation will include completion of two separate surveys. The initial survey is to be completed after consent is obtained and a 30 day follow-up survey will be sent to the patient. Each of the two surveys is similar in structure and each take 5 to 7 minutes to complete. The follow-up survey will be administered via the U.S. Postal Service mail. The patient will be asked to complete the survey and return it in a self addressed, postage paid envelope provided. Neither survey will contain any patient specific identifiable information.
I hereby acknowledge that ______________________ has provided information about the procedure described above, about my rights as a subject, and he answered all questions to my satisfaction. I understand that I may contact him at Phone No. (614) 292-1716 should I have additional questions. He has explained the risks described above and I understand them; he has also offered to explain all possible risks or complications.

I understand that, where appropriate, the U.S. Food and Drug Administration may inspect records pertaining to this study. I understand further that records obtained during my participation in this study that may contain my name or other personal identifiers may be made available to the sponsor of this study. Beyond this, I understand that my participation will remain confidential.

I understand that I am free to withdraw my consent and participation in this project at any time after notifying the project director without prejudicing future care. No guarantee has been given to me concerning this treatment or procedure.

I understand in signing this form that, beyond giving consent, I am not waiving any legal rights that I might otherwise have, and I am nor releasing the investigator, the sponsor, the institution, or its agents from any legal liability for damages that they might otherwise have.

In the event of injury resulting from participation in this study, I also understand that immediate medical treatment is available at University Hospitals of the Ohio State University and that the costs of such treatment will be at my expense; financial compensation beyond that required by law is not available. Questions about this should be directed to the Office of Research Risks at 292-5958.

I have read and fully understand the consent form. I sign it freely and voluntarily. A copy has been given to me.

Date _______ Time _____ am/pm Signed __________________________ (subject)

Witness(es) _____________________________ (Person Authorized to Consent for Subject if required)

if required

I certify that I have personally completed all blanks in this form and explained them to the subject or his/her representative before requesting the subject or his/her representative to sign it.

Date ___________ Signed ____________________________

(Signature of Project Director or his Authorized Representative)

HS-028A (Rev. 7/93)
APPENDIX F

Baseline Health Status Survey

256
Cost, Outcomes and Physician Compliance with Clinical Practice Guidelines: A Case Study of Vancomycin Use and Intravenous to Oral Antimicrobial Guidelines at The Ohio State University Medical Center

THE OHIO STATE UNIVERSITY

Code No.  

Protocol No. 99HO393

257
Instructions: This questionnaire asks for your views about your health. This information will help keep track of how you feel and how well you were able to do your usual activities just after you were recently discharged from the hospital (in other words, the first week you were home from the hospital).

Please answer every question by marking one box. If you are unsure about how to answer, please give the best answer you can.

1. In general, would you say your health was:
   0 Excellent 0 Very good 0 Good 0 Fair 0 Poor

The following items are about activities you might do during a typical day. Did your health limit you in these activities? If so, how much?

<table>
<thead>
<tr>
<th></th>
<th>Yes, Limited</th>
<th>Yes, Limited</th>
<th>No, Not Limited</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A Lot</td>
<td>A Little</td>
<td>At All</td>
</tr>
</tbody>
</table>

2. **Moderate activities**, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf
   O O O

3. Climbing several flights of stairs
   O O O

During the week after discharge, did you have any of the following problems with your work or other regular daily activities as a result of your physical health?

<table>
<thead>
<tr>
<th></th>
<th>YES</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Accomplished less than you would like
   O O

5. Were limited in the kind of work or other activities
   O O

During the week after discharge, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

<table>
<thead>
<tr>
<th></th>
<th>YES</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. Accomplished less than you would like
   O O

7. Didn’t do work or other activities as carefully as usual
   O O

8. During the week after discharge, how much did pain interfere with your normal work (including both work outside the home and housework)?
   O O O O O
   Not at all A little bit Moderately Quite a bit Extremely

258
These questions are about how you felt and how things were with you during the week after discharge. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the week after discharge:

<table>
<thead>
<tr>
<th>Question</th>
<th>Time</th>
<th>All</th>
<th>Most</th>
<th>A Good</th>
<th>Some</th>
<th>A Little</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>9. Did you felt calm and peaceful?</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10. Did you have a lot energy?</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11. Did you feel downhearted and blue?</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

12. During the week after discharge, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

<table>
<thead>
<tr>
<th>Time</th>
<th>All</th>
<th>Most</th>
<th>A Good</th>
<th>Some</th>
<th>A Little</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The following questions refer to your health experiences the over the 30 days prior to being admitted to the hospital. During that 30 days:

13. How many days were you unable to perform your normal daily activities due to illness, prior to being admitted to the hospital?

   _______ (Number of days between 0 and 30)

14. How many times did you visit an Emergency Department for illness in the 30 days prior to being admitted to the hospital?

   _______ (Number of visits)

15. Before being admitted to this hospital this last time, were you admitted to this hospital or another hospital within the past 30 days?

   O  YES  O  NO

   If yes, were you treated for an infection?

   O  YES  O  NO

16. Prior to being in the hospital, how many physician office visits did you make within the prior 30 days?

   _______ (Number of physician office visits)

   259
17. Please mark the highest level of education you have obtained:

0 some high school
0 high school graduate
0 some college
0 college graduate
0 advanced college degree

18. Prior to being admitted to the hospital, were you:

0 working full time
0 working part time
0 not working
0 a full time student
0 retired

THANK YOU FOR PARTICIPATING IN THIS SURVEY!
APPENDIX G

Chart Review Form
Data Collection Form

Medical Chart Review

Code #: ____________________________ Date: ____________________________

I. Demographics

A. Name: ____________________________

B. Medical Record Number: ______________

A. Date of Birth __/__/____

B. Age: _______

C. Gender M F Race: _______

D. Admit Service
   _____ Medical Service
   _____ Surgical Service
   _____ To ICU
   _____ Other

E. Attending MD: ______________________

F. Primary Payer
   _____ Medicaid
   _____ Medicare
   _____ Private - Managed Care
   _____ Private – Other
   _____ Public – Other
   _____ Self-Pay
   _____ No Pay
   _____ Second

G. Marital Status
   _____ Married
   _____ Single

H. Occupation: _________________________

I. Billing DRG: ________________________
J. DRG Codes

1. ________
2. ________
3. ________
4. ________
5. ________
6. ________
7. ________
8. ________
9. ________
10. ________
11. ________
12. ________
13. ________
14. ________
15. ________
16. ________
17. ________
18. ________
19. ________
20. ________
21. ________
22. ________
23. ________
24. ________
25. ________
26. ________
27. ________
28. ________
29. ________
30. ________

II. Clinical History

A. Chief Complaint: ________________________________

B. Comorbid Conditions

<table>
<thead>
<tr>
<th>Chronic Conditions</th>
<th>AIDS</th>
<th>active cancer</th>
<th>positive cancer history</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>____</td>
<td>____</td>
<td>____</td>
</tr>
<tr>
<td></td>
<td>____</td>
<td>____ diabetes</td>
<td>____ acute ____</td>
</tr>
<tr>
<td></td>
<td>____</td>
<td>____ renal disease</td>
<td>____</td>
</tr>
<tr>
<td></td>
<td>____</td>
<td>____ organ transplant</td>
<td></td>
</tr>
<tr>
<td></td>
<td>____</td>
<td>____ cardiovascular disease</td>
<td>____ MI ____ CHF</td>
</tr>
<tr>
<td></td>
<td>____</td>
<td>____ pulmonary disease</td>
<td>____ asthma ____ COPD</td>
</tr>
<tr>
<td></td>
<td>____</td>
<td>____ sepsis</td>
<td>____ other ____</td>
</tr>
<tr>
<td></td>
<td>____</td>
<td>____ other</td>
<td>____ other ____</td>
</tr>
<tr>
<td></td>
<td>____</td>
<td>____ other</td>
<td>____ other ____</td>
</tr>
</tbody>
</table>
C. Drug Therapy On Admission

<table>
<thead>
<tr>
<th>Drug</th>
<th>Strength</th>
<th>Regimen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

F. Antibiotic Drug Therapy upon admission  yes  no

III. Guideline Compliance

A. IV to PO switch

1. Date First PO Medication  __/__/__
2. Date First PO Nutrition  __/__/__
3. Date IV switched to PO  __/__/__
4. Additional Antimicrobial Therapy during hospital admission:

<table>
<thead>
<tr>
<th>Drug</th>
<th>Strength</th>
<th>Regimen</th>
<th>Date - Started</th>
<th>Date - Discontinued</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. Did the patient have either of the following diagnoses:

- Endocarditis: yes, no
- Meningitis: yes, no

B. IV Vancomycin

1. Beta Lactam resistant gram + organism culture

   - yes, no

2. Number and result

   - First Date: / / , Result: + -
   - Second Date: / / , Result: + -
   - Third Date: / / , Result: + -

3. Documented Beta Lactam Allergy

   - yes, no
4. Purpose of Therapy
   - prophylactic
   - empiric treatment
   - confirmed lab treatment

5. Diagnosis of Endocarditis
   - yes
   - no

6. Did the patient have thoracic surgery
   - yes
   - no

IV. Outcome Data

A. Infection Symptom Resolution

1. Temperature > 101.5 within the first 24 hours after enrollment (IV therapy).
   - Yes
   - No
   - Number of consecutive days temp > 101.5 after starting antimicrobial therapy

2. WBC count > 10 within the first 24 hours after enrollment (IV therapy)
   - Yes
   - No
   - Number of consecutive days WBC > 18,000 after starting antimicrobial therapy

3. Pulse Rate
   - Yes
   - No
   - Number of consecutive days WBC > 18,000 after starting antimicrobial therapy

4. Success
   - Failure

266
## B. Discharge Medication

<table>
<thead>
<tr>
<th>Drug</th>
<th>Strength</th>
<th>Regimen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX H

Follow-Up Health Status Survey
Cost, Outcomes and Physician Compliance with Clinical Practice Guidelines: A Case Study of Vancomycin Use and Intravenous to Oral Antimicrobial Guidelines at The Ohio State University Medical Center

THE OHIO STATE UNIVERSITY

Code No. _______________
Protocol No. 99HO393

269
Instructions: This questionnaire asks for your views about your health. This information will help keep track of how you feel and how well you are able to do your usual activities.

Please answer every question by marking one box. If you are unsure about how to answer, please give the best answer you can.

1. In general, would you say your health was:

   Excellent  Very good  Good  Fair  Poor

   O  O  O  O  O

The following items are about activities you might do during a typical day. Did your health limit you in these activities? If so, how much?

2. Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf

   Yes, Limited  Yes, Limited  No, Not
   A Lot  A Little  At All

   O  O  O

3. Climbing several flights of stairs

   O  O  O

During the past week, did you have any of the following problems with your work or other regular daily activities as a result of your physical health?

4. Accomplished less than you would like

   YES  NO

   O  O

5. Were limited in the kind of work or other activities

   O  O

During the past week, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

6. Accomplished less than you would like

   YES  NO

   O  O

7. Didn’t do work or other activities as carefully as usual

   O  O

8. During the past week, how much did pain interfere with your normal work (including both work outside the home and housework)?

   Not at all  A little bit  Moderately  Quite a bit  Extremely

   O  O  O  O  O

270
These questions are about how you felt and how things were with you during the past week. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the past week -

<table>
<thead>
<tr>
<th>Question</th>
<th>All of the Time</th>
<th>Most of the Time</th>
<th>A Good Bit of the Time</th>
<th>A Little of the Time</th>
<th>None of the Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>9. Did you feel calm and peaceful?</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>10. Did you have a lot of energy?</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>11. Did you feel downhearted and blue?</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

12. During the past week, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

<table>
<thead>
<tr>
<th>Time</th>
<th>All of the time</th>
<th>Most of the time</th>
<th>Some of the time</th>
<th>A little of the time</th>
<th>None of the time</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
</tbody>
</table>

The following questions refer to your health experiences over the past 30 days after being discharged from The Ohio State University Medical Center. During the 30 days-

13. Since being discharged, approximately how many days were you unable to perform your normal daily activities as a result of infection?

_____ (Number of days between 0 and 30)

14. Since being discharged, how many times did you visit an Emergency Department for treatment of an infection?

_____ (Number of times visited an Emergency Department)

15. Were you admitted to a hospital again for the treatment of infection?

O YES O NO

16. Since being discharged, how many physician office visits did you make for the treatment of infection?

_____ (Number of physician office visits)

17. If you visited a physician, were you treated with antibiotic therapy?

O YES O NO
18. Since you were last admitted to the hospital, has there been any change in your overall quality of life:
   O yes, improvement in overall quality of life (go to question 19 and stop)
   O no change in quality of life (you are finished)
   O yes, deterioration in overall quality of life (go to question 20 and stop)

19. If you responded yes to question 18 that your overall quality of life has improved, please rate how much your overall quality of life has changed:
   O Almost the same, hardly any better at all
   O A little better
   O Somewhat better
   O Moderately better
   O A good deal better
   O A great deal better
   O A very great deal better

20. If you responded yes to question 18 that your overall quality of life has deteriorated, please rate how much your overall quality of life has changed:
   O Almost the same, hardly any worse at all
   O A little worse
   O Somewhat worse
   O Moderately worse
   O A good deal worse
   O A great deal worse
   O A very great deal worse

THANK YOU FOR PARTICIPATING IN THIS SURVEY!
APPENDIX I

Investigational Review Board Approval
With regard to the employment of human subjects in the proposed research:

99H0393 COST, OUTCOMES AND PHYSICIAN COMPLIANCE WITH CLINICAL PRACTICE GUIDELINES: A CASE STUDY OF VANCOMYCIN USE AND INTRAVENOUS TO ORAL ANTIMICROBIAL GUIDELINES AT THE OHIO STATE UNIVERSITY MEDICAL CENTER, Dev S. Pathak, Dennis W. Grauer, Mitchell S. Medow, Center for Health Outcomes, Policy, and Evaluation Studies (HOPES)

THE BIOMEDICAL SCIENCES REVIEW BOARD HAS TAKEN THE FOLLOWING ACTION:

X APPROVED WITH STIPULATION(S)* ___ WAIVER OF WRITTEN CONSENT GRANTED

*Stipulation(s) stated by the IRB have been met by the investigator, and therefore, the protocol is APPROVED.

It is the responsibility of the principal investigator to retain a copy of each signed consent form for at least three (3) years beyond the termination of the subject's participation in the proposed activity. Should the principal investigator leave the University, signed consent forms are to be transferred to the Human Subjects Review Board for the required retention period. This application has been approved for a period of not more than one year. You are reminded that you must promptly report any problems to the Review Board, and that no procedural changes may be made without prior review and approval. You are also reminded that the identity of the research participants must be kept confidential.

Date: December 13, 1999

Signed: ____________________________
Chairperson

HS-025H (Rev. 9/99).
APPENDIX J

Guideline Criteria Development Tool
## Standards of Guideline Development

<table>
<thead>
<tr>
<th></th>
<th>Converting Intravenous to Oral Antimicrobials</th>
<th>Criteria for Use of Vancomycin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Purpose of the guideline is specified</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Rationale and importance of the guideline are explained</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>The participants in the guideline development process and their areas of expertise are specified</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Targeted health problem or technology is clearly defined</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Targeted patient population is specified</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Intended audience or users of the guideline are specified</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>The principal preventive, diagnostic, or therapeutic options available to clinicians and patients are specified</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>The health outcomes are specified</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>The method by which the guideline underwent external review is specified</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>An expiration date or date of scheduled review is specified</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Method of identifying scientific evidence is specified</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Time period from which evidence is reviewed is specified</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>The evidence used is identified by citation and reference</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Method of data extraction is specified</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Method for grading or classifying the scientific evidence is specified</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Formal methods of combining evidence or expert opinion are used and described</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Benefits and harms of specific health practices are specified</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Benefits and harms are quantified</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>The effect on health care costs from specific health practices is specified</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Costs are quantified</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>The role of value judgments used by the guideline developers in making recommendations is discussed</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>The role of patient preferences is discussed</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Recommendations are specific and apply to the stated goals of the guideline</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Recommendations are graded according to the strength of the evidence</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Flexibility in the recommendations is specified</td>
<td></td>
</tr>
</tbody>
</table>

**TOTAL STANDARDS ENDORSED:** 276
APPENDIX K

Follow-Up Cover Letter

277
March 10, 2000

I would like to thank you again for participating in this study. We are conducting a short, mailed survey research project to measure how well patients recently treated at The Ohio State University Medical Center feel. This second and final survey is part of a research project designed to examine the decisions made with respect to health care services provided and the impact of these decisions on our patients. This project will help measure how you are currently feeling.

Please note that participation in the study is confidential and voluntary. All information obtained from the study is kept strictly confidential. No information shared by any patient will be given to your physician or anyone else. All data will be reported as group (over all participants) information. Thus, the information will not have any impact on current or future care obtained at Ohio State University Medical Center.

We have made significant effort to keep the survey short (less than five minutes) to reduce the amount of your time involved in the study. For the second follow-up survey, participation requires the following:

5. Complete the follow-up survey.
6. Return the survey in the postage paid envelope.

This is the final survey to be sent to you. It is extremely important that each participant who completed the initial survey also completes the second, follow up survey. The initial survey asked about how you felt after returning from the hospital, while this second survey will ask about how you currently feel. The two surveys allow us to measure if you are feeling better or not. To participate in the research it will take approximately 10 minutes of your time (5 minutes for each survey).

Your participation in the study is greatly appreciated! If you have any questions or concerns with participating in the study, do not hesitate to call me at 614 292-1716. This research project is part of my dissertation at Ohio State to complete my Ph.D. I have always been interested in studying if patients truly feel better after the care they received. Thus I have an interest in measuring the way you 'feel'. Again, if you have any questions, please give me a call or send me a note.

Sincere Thanks,

Dennis W. Grauer, M.S.
Phone: 614-292-1716

278
BIBLIOGRAPHY


Bulpitt CJ. Randomized Controlled Clinical Trials. The Hague, the Netherlands: Martinus Nijhoff;1983.


288


Hayward RSA, Guyett GH, Moore KA, McKibbon KA, Carter AO. Canadian physicians' attitudes about the preference regarding clinical practice guidelines. *CMAJ.* 1997;156(12):1715-1723.


James BC. Implementing practice guidelines through clinical quality improvement. Front Health Serv Manage. 1993;10(1):54-56.


294


296

Lembcke PA. Medical auditing by scientific methods. JAMA. 1956;162:646-655.


297


Speither CE. Practice parameters. An opportunity for pathologists to take a leadership role in patient care. *Arch Pathol Lab Med.* 1990;114(8):823-824.


*SPSS® Base 9.0 Applications Guide.* Chicago, IL;SPSS, Inc;1999.


Torrance ME. Understanding Epidemiology. St Louis, MO; Mosby-Year Book, Inc; 1997.


Unpublished Data. The Ohio State University Medical Center. Columbus, OH; Department of Pharmacy, 1999.

Unpublished Data. The Ohio State University Medical Center. Columbus, OH; The Ohio State University: 1997.


Ware JE, Kosinski M, Keller SD. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales. Boston, MA; The Health Institute, New England Medical Center: 1995.


