AN APPLICATION OF MULTI-ATTRIBUTE UTILITY THEORY TO MEDICAL REFERRAL DECISIONS IN THE MANAGEMENT OF OBESITY

By

Penny Annette Jennett

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Administration and Curriculum

1982

ABSTRACT

AN APPLICATION OF MULTI-ATTRIBUTE UTILITY THEORY TO MEDICAL REFERRAL DECISIONS IN THE MANAGEMENT OF OBESITY

By

Penny Annette Jennett

Good clinical judgment is of cardinal importance to the medical profession. The uncertainty and complexity of the clinical environment and the limitations and biases of human information processing can hinder optimum decision making. Decision analysis can be particularly useful for analyzing complex clinical problems because it is devised to compensate for these limitations and biases. Multi-attribute utility theory (MAUT) is one of a number of available decision aids. Its orientation and focus upon patient outcomes may be particularly compatible with the clinical environment.

This study examined how the MAUT model performed when applied to the long-term ambulatory problem of obesity. The model was used to analyze the management decisions of a sample of 45 primary care physicians to refer or not to refer obese female patients to an endocrinologist. Study subjects were primary care physicians within the specialties of Internal Medicine, Family Practice, and Gynecology. The decisions of the subjects generated from the model were compared to their referral decisions made to a series of case vignettes depicting similar patients.

Six research questions were asked in the study. The data collected were physician responses to a series of case vignettes and to a semi-structured interview. Linear regressions, one-way ANCOVAs and MANCOVAs, one-way ANOVAs, Chi-Squares, and descriptive techniques were used to analyze the data.

Any differences reported to be significant were assessed at the .05 alpha level. Two major findings are reported:

1. The predictors of physician referral behavior derived from the model were not useful measures for predicting physician vignette referral behavior.

2. The predictors of physician referral behavior derived from the model did not vary by selected physician characteristics.

Interpretation of the findings demonstrates that MAUT applications can show decision makers if a decision is or is not sensitive to factor(s) they previously thought were important. The model also permits decision makers to determine if their choices are consistent with the stated preferences and uncertainties of a given situation. These opportunities have important implications for research, decision analysis, quality patient care, medical education, and health policy.

ACKNOWLEDGEMENTS

I acknowledge the special assistance and support that individuals have offered me throughout my doctoral program. I aspire to be worthy of these efforts as I continue with my life and career.

Dr. Richard Featherstone, Chairperson of my Doctoral Committee, has my admiration and appreciation for his outstanding academic and personal integrity and for the superior role he played as my Academic Advisor. Special thanks are directed to him for his guidance throughout my academic coursework and preparation for the written doctoral comprehensive examinations. His unfailing support, respect, sincerity, and friendship are deeply appreciated.

I want to express my gratitude to Dr. Arthur Elstein for the learning opportunities and academic guidance he has offered. I am indebted to him for providing the research setting for my dissertation, for agreeing to direct this dissertation, and for giving the time and commitment this responsibility has required. He has also given me the opportunity to share the knowledge and skills of his entire clinical decision making research team. His scholastic endeavors, creative thinking, precision in verbal and written communication, and rigor in research serve as models for my future academic pursuits.

Dr. Howard Teitelbaum and Dr. Larry Lezotte have my appreciation for agreeing to act as members of my Doctoral Committee and for their subsequent help and encouragement. Special thanks is given to Dr. Teitelbaum for his conscientious guidance as my Faculty Advisor during my Masters studies.

Interaction with Dr. Elstein's decision making research team has had an enormous impact on my academic growth. I acknowledge with deepest appreciation each of the member's contributions to my graduate education. In particular, I am grateful to Dr. Marilyn Rothert for her dedicated efforts as Program Associate in the coordination, organization, and design of the study upon which my dissertation is based.

I am indebted to Nova Green for her assistance in the editing and final preparation of my dissertation and for helping to make this endeavor a scholastically rewarding experience.

I wish to thank Necia Black and Bill Metheny for their conscientious and professional approach to data collection which led to the data base upon which my dissertation is based. Necia Black acted as my Office of Research consultant and provided invaluable assistance and service in data entry and analysis.

Patricia Borowiak and Judy Zivick provided outstanding professional secretarial and typing service throughout my graduate studies. I am deeply indebted to them both for the high quality of their work, their patience and good humor with unreasonable deadlines and copious drafts, and their personal sacrifices to help me fulfill my commitments. Pat Borowiak's secretarial and transcriptional support through my academic coursework could not be surpassed. Judy Zivick contributed an equalled excellence and personal interest to my program this year. She merits my deepest thanks for taking the responsibility of typing the final draft of my dissertation.

I am grateful to the Office of Medical Education Research and Development at Michigan State University for the opportunity to participate as an OMERAD Fellow for the entire period of my Masters and Ph.D. graduate studies. The support I have received from the Office has greatly facilitated my ability to achieve my academic goals. The secretarial support service I have been accorded over the past two months deserves special mention.

The long-term respect and support that I have received from Dr. O. E. Laxdal have given me the courage to pursue and achieve this goal.

I am deeply appreciative for the contributions and love of my family and friends.

The research was supported in part by a grant from the National Library of Medicine (NLM), LM-03396.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND STATEMENT OF THE PROBLEM

Introduction

"One of the arts for the decision analyst is the art of knowing how much and what kind of decision analysis to do. The degree of analysis can range from making simple lists to constructing giant interactive computer models. To be effective, decision analysis must be appropriate: the extent of the analysis must be suitable to the means and ends of the decision maker. The question of whether the analysis was appropriate to the decision maker and his problem is one that should always be raised in judging effectiveness." (Howard, 1980)

Judgment is and always has been cardinal to medicine (Price et al., 1971). Physicians, as clinical decision makers, value accuracy in their professional judgments (Ginsburg and Offensend, 1968). Indeed, good clinical judgment has been rated as the foremost attribute desired in physicians (Price et al., 1971). Patient management decisions require clinicians continually to make complex and critical diagnostic and therapeutic choices based upon what they expect and prefer to happen. These predictive and evaluative clinical judgments are for the most part intuitive, i.e., made without formal explication of the reasoning process.

While the medical profession desires accurate judgments, physicians, as human problem solvers within complex clinical environments, are influenced by a number of variables that may cause suboptimal judgments to be made. Research in cognitive processes has repeatedly verified that human problem solvers must omit certain features of complex environments because their information processing takes place in a problem space of limited size (Newell and Simon, 1972). In addition, decision makers are predisposed to biases which can influence their decisions (Newell and Simon, 1972; Nisbett and Ross, 1980; Slovic et al., 1982). Problem solvers' subjective assessments about the likelihood of certain events and their perception of the correlation between events are influenced by their prior experiences and expectations and by the ease with which similar instances can be brought to mind. These beliefs tend not to be revised on the basis of evidence (Slovic and Lichtenstein, 1971; Tversky and Kahneman, 1974). There is a tendency for decision makers to be affected by the manner in which the problem is structured, to overestimate small probabilities and underestimate large probabilities, to display shortcomings in their ability to weight and combine information, and to be insensitive to sample size (Nisbett and Ross, 1980). As well, attitudes based on the social and economic status of clinicians and/or patients, time constraints, and the available health care system can influence the definition of the problem and affect the implementation of health care management decisions (Eisenberg, 1979).

The complex nature and potentially critical consequences of patient management decisions as well as the medical profession's concern for precision and accuracy have stimulated the study of methods to facilitate optimal clinical decisions. Hammond and Joyce (1977) suggest two ways to increase the precision of medical judgments: 1) increase medical knowledge about the problem in question, and 2) improve consistency in information processing.

The application of decision analysis to clinical problems can help the medical profession with the second task. Decision analysis is the systematic application of decision models to problem situations to permit the evaluation of alternative actions. It requires the explicit separation of uncertain and complex decision problems into manageable parts followed by the recombining of these parts according to normative rules. Decision analysis is particularly useful for analyzing complex clinical problems because it is devised to compensate for the clinical problem solvers' limitations and biases by assisting in the apprehension, selection, and combination of complex multiple clues. It facilitates incorporating relevant values and uncertainties in a systematic and unbiased manner. Improvement in the quality of patient care has been reported as the result of the use of decision analysis (Sisson et al., 1976). This study considers the application of a particular model, Multi-Attribute Utility Theory, to the decision analysis of a clinical problem.

Multi-Attribute Utility Theory

Any decision problem is characterized by the availability of more than one alternative course of action. Multi-Attribute Utility Theory (MAUT) is a system for assessing the outcomes of such alternative courses of action (Huber, 1974). In complex decision situations, an outcome resulting from a chosen alternative rarely can be described completely in terms of only one attribute. Outcomes described by more than one attribute are referred to as multi-attributed outcomes. Evaluating a multi-attributed outcome is difficult because the exact contribution of each attribute to a given outcome is uncertain, and decision makers differ in their preferences for various outcomes. MAUT models are designed to assess the combined effect of each attribute in terms of the avowed preferences of the decision maker.

To assess management alternatives by the use of MAUT, one proceeds through the following steps (Johnson and Huber, 1977; Edwards, 1977):

1. The decision problem is structured. This step involves identifying the key decision makers, the decision alternatives, and the significant attributes which may have values in describing the outcome of a particular choice of action.

2. The decision makers' preferences (utilities) are assessed for each of the identified attributes.

3. The uncertainties (probabilities) are assessed for each of the identified attributes.

4. The alternatives are evaluated. The numerical values obtained for each attribute in steps 2 and 3 are aggregated across attributes using a suitable aggregate and weighting scale. This mathematical translation results in a predictor score called a subjective expected utility (SEU) for each alternative course of action. This score can act as an aid to determine the choice of management alternatives.

5. An alternative is chosen. The alternative with the highest SEU is considered to be the best balance between risks and benefits.

For example, in the medical or surgical management of coronary heart disease, each possible patient outcome state is discussed in terms of two attributes: quality of life (the degree of relief from chest pain) and survival (longevity). The attribute of degree of relief from chest pain can be further divided into two levels: chest pain relieved (c+) and chest pain not relieved (c-). The survival attribute can be considered from three levels: long-term survival (at least 10 years) (s+), short-term (less than 10 years) (s-), and death within 30 days of surgery (d) (Weinstein et al., 1980). A patient outcome state could consist of any of the following combinations: c+,s+; c+,s-; c-,s+; c-,s-; and d. Each of these combinations has specific probabilities depending upon which management alternative (i.e., medical or surgical) is chosen, and preferences depending upon the decision maker. By analyzing the probabilities and preferences for each of the attributes within the stated patient outcomes, the possible patient consequences of choosing either management alternative can be more clearly and directly assessed.

Acceptability of the Model

Simple MAUT applications may be particularly acceptable to the clinical world because their focus is on patient outcomes (Williamson et al., 1975). In choosing management strategies, physicians commonly agree that there should be a positive outcome for the patient from the selected treatment. The use of outcomes as criteria for assessing quality of care is recommended by a number of clinicians. Laxdal et al. (1978) suggest using measures of preventable morbidity, mortality, cost, and patient satisfaction in quality care programs, and report that physicians are motivated to participate in programs where the procedures are clearly related to measures of improvement in patient care and health status outcomes. MAUT applications also combine the probabilistic character of the medical environment with current knowledge, incorporate a way of balancing the benefits and risks to the patient of certain procedures, and explicitly outline the expected utility of various alternatives when considering choices in diagnosis and therapy. Childs and Hunter (1972) and Donabedian (1976) report these are key considerations that characterize the practice of a good physician and are at the heart of technical quality.

MAUT has been described as a method for dealing with life threatening clinical problems such as renal failure (Gorry et al., 1973), cancer (McNeil and Pauker, 1979), and burns (Gustafson and Holloway, 1975). Models could also be used to formulate decision rules to clarify difficult management decisions frequently encountered in office or ambulatory clinic-based practice. Common long-term ambulatory problems such as hypertension, obesity, and diabetes mellitus present complex management choices. In the literature reviewed, however, no applications of a MAUT model to long-term ambulatory problems were found. Johnson and Huber (1977) and Howard (1980) recommend research be carried out to discover how specific analysis models will perform when applied to different types of clinical problems. This study examined how the MAUT model performed when applied to the specific long-term ambulatory problem of obesity.

Obesity

Obesity causes or exacerbates several health problems. Van Itallie (1979) notes that the risk to health increases with the percentage of overweight, and that weight reduction improves the patient's medical, physical, social, and psychological health. Obesity is a prevalent long-term problem in America. The National Center for Health Statistics data for 1971-74 (Abraham and Johnson, 1979) report that 14% (8,041) of men age 20-74 and 23.8% (15,268) of women age 20-74 are 20% or more above desirable weight. Desirable weight is defined as the mean weights for men and women aged 20-29 years. The obese middle-aged female population is most at risk for morbidity and mortality (Van Itallie, 1979). The physician is faced with the problem of helping obese patients lose weight and maintain the loss over time. Several management strategies are available that may or may not achieve this goal. Medical, surgical, psychotherapeutic, and dietary methods of weight reduction exist.

A chart study by Ravitch et al. (1982) demonstrated that primary care physicians refer obese patients who do not have classic signs or symptoms of endocrine-based obesity to endocrinologists. As this referral rate is greater than the incidence of obesity due to endocrine or metabolic disorder, and the referral of patients to sub-specialists generates health costs, it seemed worthwhile to investigate the reasons behind these decisions. Therefore, for this research, two treatment strategies used by primary care physicians in the management of obese female patients were framed by the researchers. These strategies were referral to an endocrinologist and no referral. It was assumed that the significant attributes of the outcomes resulting from the management alternatives of referral or non-referral were weight reduction and patient satisfaction with management.

Purpose and Research Questions

The purpose of this dissertation was to examine how the MAUT model performed when applied to the long-term ambulatory problem of obesity. The two attributes chosen to characterize the health care outcomes within the management alternatives of patient referral or non-referral to an endocrinologist were weight reduction and patient satisfaction with management. Specifically, the model was applied to the referral management decisions of primary care physicians dealing with two cohorts of 50 percent and 100 percent overweight middle-aged female patients. The decisions derived from the model were compared to intuitive management decisions made to a series of case vignettes depicting similar patients. The study subjects were 45 voluntary primary care physicians with a varied number of years of practice experience, within the specialties of Internal Medicine, Family Practice, and Gynecology. Data were collected from physician responses to a series of case vignettes and to a semi-structured interview. The following research questions were considered in the data collection.

A. How closely do the predictors of physician referral behavior derived from the model relate to physicians' management decisions in response to the case vignettes?

1. What is the relation between the predictors of physician referral behavior derived from the model and number of vignette cases physicians referred? Is it possible to predict the number of vignette cases referred from scores derived from the model?
2. What is the relation between the predictors of physician referral behavior derived from the model and number of vignette cases physicians referred <u>for each specialty</u>? Does the MAUT model fit one specialty group better than another?
3. What is the relation between the predictors of physician referral behavior derived <u>from each attribute</u> and number of vignette cases

physicians referred? Is it possible to predict the number of vignette cases referred from the scores derived for each attribute? Of the two attributes, weight reduction and patient satisfaction, is one attribute more highly predictive of number of vignette cases referred than the other?

B. What effect do selected physician characteristics and categories have on these predictors of physician referral behavior derived from the model?

4. Do the predictors of physician referral behavior derived from the model vary by physician characteristics? The characteristics analyzed were specialty and years in practice.

5. What are the properties of the individual attribute parameters (probabilities, importance weights, and values)? Are individual attribute parameters affected by selected physician characteristics (specialty, years in practice) or weight categories (50% and 100% overweight)? What is the relation between individual attribute parameters and number of vignette cases referred?

6. Do the preferred management alternatives, to refer or not to refer, vary by selected categories? The categories examined were sex, degree (D.O., M.D.), practice type (Academic, Community), practice location (Lansing, Grand Rapids), and perception of self as overweight.

Importance of Study

This study builds on work by researchers who have investigated the use of decision analysis in medical treatment management decisions. Decision analysis has not been applied extensively in the health field (Weinstein et al., 1980). Studies such as this may provide greater insight into how such models can be used in analyzing clinical problems and lead to the development of operational models that can clarify clinical management decisions and serve as educational tools.

As treatments become increasingly expensive and potentially invasive, clinical decisions require additional analytical validation (Knowles, 1977). Applications of decision analysis provide a mechanism for validation.

Value judgments underlie virtually all clinical decisions. Therefore, it is important that physicians understand what outcomes they value and why (Weinstein et al., 1980). Shepard (1964) points out that people may use quite different evaluation models than the ones they believe they are using. A formalized decision analysis approach that makes the decision process explicit can help decision makers be aware of the components of their decision.

Both patients and society have demanded an increased participation in clinical decision making. Decision analysis techniques can help all parties involved in the decisions to recognize and reconcile their different evaluations of the probabilities, risks, and benefits. Alternatives in treatment strategies may be suggested when the values and priorities of all parties are clearly understood (McNeil et al., 1978; McNeil and Pauker, 1979).

Obesity, the condition to which decision analysis is applied, is a prevalent long-term problem in America. Several modes of therapy are available for the treatment of the condition. Successful treatment of obesity has distinct health benefits, but there is a great deal of uncertainty among practitioners and patients regarding the selection of the most effective treatment so that beneficial outcomes occur and are maintained (Maddox et al., 1968). If decision analysis can be shown to clarify values and actions and to simplify management choices when applied to this particular condition, it would be an important step toward improving clinical care for the obese patient.

Interpretation of data obtained from this study should increase understanding of the appropriateness of other clinical applications of multi-attribute utility theory. Information resulting from the analysis of the probabilities and preferences that physicians place on attributes of potential patient outcomes in the problem situation of obesity could lead to important insights into the usefulness of MAUT clinical applications to other long-term ambulatory problems.

Assumptions

This dissertation is based on three assumptions:

1. Good clinical judgment is of cardinal importance to the medical profession.

2. The uncertainty and complexity of the clinical environment and the limitations and biases of human information processing can hinder optimum decision making. Decision analysis can be particularly useful for analyzing complex clinical problems because it is devised to compensate for these limitations and biases.

3. Multi-attribute utility theory, because of its focus upon patient outcomes, may be particularly compatible with the values of clinicians.

Limitations

This dissertation does not examine or address the following issues:

1. Which management alternative, "refer" or "not refer," is better for obesity, i.e., if patients do, in fact, lose more weight or are more satisfied with referral or non-referral.

2. The causes of obesity.

3. Management strategies other than referral or non-referral to an endocrinologist.

4. Attributes other than weight reduction and patient satisfaction with management.

5. The monetary costs of either alternative.

6. Values or opinions other than those of the physician subjects.

CHAPTER II

REVIEW OF LITERATURE ON DECISION MAKING THEORY, MODELS, AND RESEARCH

Decision Analysis Theory and Models

The concepts and early applications of decision analysis originated out of economics, marketing, military planning, and applied mathematics (Kassirer, 1976). Persons working in these complex environments recognized the need for rules and models to simplify and clarify decision making. Examples of techniques developed to aid decision making range from agenda lists, flow charts, and algorithms, to complex decision analytic models, including multi-attribute utility theory.

Decision analysis is an approach to decision making under conditions of uncertainty (Howard, 1968). It involves the systematic application of decision models to problem situations to permit the evaluation of alternative actions. It requires separating uncertain and complex decision problems into manageable parts and then recombining these parts according to normative rules. Decision analysis is centrally concerned with what should be done and decision models are, therefore, prescriptive. They "do not just describe, but aid decision makers in deciding what they should do...so that their decisions will be consistent with their underlying assessments of the problem, of the uncertainties, and of the valued outcomes." (Weinstein et al., 1980)

Decision analysis is particularly useful for analyzing problems in complex environments because it is devised to compensate for limitations of human problem solving which affect intuitive judgment. These limitations and biases have been documented through extensive research (Newell and Simon, 1972; Slovic and Lichtenstein, 1971; Tversky and Kahneman, 1974; Nisbett and Ross, 1980; Hogarth, 1980). It has been shown that the finite capacity of the decision maker's short-term memory produces the phenomenon of "bounded rationality," a defined area of time or space within which a problem solution is sought. Research in cognitive processes has verified repeatedly that human problem solvers omit certain features of complex environments because their information processing takes place in this limited problem space (Newell and Simon, 1972). The limited problem space also requires the information processor to select and process data serially. Solutions anticipated by the problem solver may bias data selection, and the sequence of data processing itself may affect the problem solution (Hogarth, 1980).

In addition to working in bounded problem spaces, decision makers are predisposed to information processing biases. Heuristics such as representativeness, availability, anchoring, and adjustment* affect probability judgments and govern estimates of subjective probability (Nisbett and Ross, 1980; Tversky and Kahneman, 1974). These heuristics may consistently distort decisions by leading to improper probability estimates. Suboptimal decisions may also result from tendencies of the decision maker to overestimate small probabilities and underestimate large probabilities, to display characteristic shortcomings in ability to weight and combine information, and to be insensitive to sample size (Nisbett and Ross, 1980).

Formal decision analysis aims to compensate for these limitations in human problem solving capability by providing a systematic method for selecting and combining complex and multiple cues, and for assessing and incorporating values

---

*Representativeness: "The degree to which the salient features of the object are representative of, or similar to, the features presumed to be characteristic of the category." (Nisbett and Ross, 1980)

Availability: "The accessibility of objects or events in the processes of perception, memory, or construction from imagination." (Nisbett and Ross, 1980)

Anchoring and Adjustment: "The failure to make necessary adjustments of initial judgments (anchors)." (Nisbett and Ross, 1980) and uncertainties. Improvement in the quality of patient care has been reported as the result of such analysis (Sisson et al., 1976).

If optimal use of decision analysis is to be made, however, the advantages and limitations of the model and its suitability for specific problem situations must be recognized and understood. Keeney and Raiffa (1972) report on these issues. Some of the advantages they identify for decision analysis are: it stimulates thinking about the various problem components, interactions between components, and relationships between problem alternatives; it promotes positive efficient interaction among key decision makers by separating the problem into component parts so that the various "experts" can address specific aspects; it provides a rationale to document the choice of a particular course of action; and, once the technique has been learned, it can be applied to similar situations. Keeney and Raiffa (1972) describe the responsibilities of the decision maker in the application of formal decision analysis. Formal decision analysis does not provide systematic procedures for isolating problems, identifying objectives, or defining measures of effectiveness for a particular situation. The decision maker is responsible for making these choices, specifying possible outcomes, outlining the attributes that define these outcomes, and specifying alternative courses of action. The decision makers must also assess uncertainties and preferences; these may be unfamiliar tasks to many. Techniques to incorporate subjective information into the model require continual revision and updating, and good outcomes are not guaranteed.

The goals of the decision makers and the constraints in their environment must be considered when evaluating the usefulness of formal decision analysis. Keeney and Raiffa (1972) emphasize that decision makers will utilize formal decision analysis only if they understand and have faith in the procedure. They must believe the analysis can help them determine their strategy and can be used to convince others to implement that strategy. Their audiences must also be able to understand the process of analysis and its results. Finally, the format chosen for analysis must not be too demanding nor time consuming.

Applicability of Decision Analysis to Clinical Problem Situations

Physicians realize that good patient management is based on more than the physician's knowledge of scientific facts. Physicians must also combine these facts appropriately, balance the dangers and discomforts of a procedure against the value of information to be gained, and recognize the probabilistic nature of patient outcomes (Donabedian, 1976). Clinicians agree that numerous factors irrelevant to considerations of the patient's medical condition are involved in management decisions, including the social and economic status of both the patient and physician, the system of care, the physician's case load, and other variables (Eisenberg, 1979; Eisenberg and Nicklin, 1981). Consequently, it is difficult to discover through questions or observations the specific criteria decision makers use in deciding between management alternatives. Also, the decision makers may believe they are evaluating alternatives on attributes quite different from those they are actually using (Shepard, 1964), or they may be uncertain which attributes they use to evaluate alternatives. They may not be aware of the attributes that apply to specific situations (Johnson and Huber, 1977). A formalized decision analysis approach to clinical problem situations can make the decision process explicit so that clinicians will have a greater awareness of the components of their decision, the outcomes they value, and why those outcomes are valued.

Expected Utility Theory

Expected Utility Theory is a major tool of decision analysis. It is a means of eliciting decision makers' preferences for certain outcomes and their estimates of the probability of these outcomes based on certain actions. These quantified outcome preferences and probabilities are combined to indicate a level of preference for certain actions or the "expected utility" of alternative actions under consideration (Weinstein et al., 1980). Daniel Bernoulli (1738) suggested the first Expected Utility model. His model was descriptive, i.e., it was intended to describe and explain the decision maker's choices. The model assumes that when any rational person is presented with two monetary choices, a gamble on a high or low return or a "sure thing" with some intermediate value, the alternative with the highest payoff will be chosen. Bernoulli devised a general method for incorporating values into expectations using a "diminishing marginal utility curve," but his model was "deficient as a guide to action for a particular individual" (Grayson, 1960). Von Neumann and Morgenstern (1953) proposed a system for determining individual utilities. They also were the first to consider the model from both the descriptive and the normative perspectives. A normative (prescriptive) model helps decision makers perceive what they should do if their decisions are to be consistent with their stated assessments of the problem, i.e., the uncertainties and the valued outcomes (Weinstein et al., 1980).

The Expected Utility model has different forms. Models can be "simple" or "complex." Simple models are those in which attributes are combined additively, while complex models use a multiplicative combining rule. There is disagreement about the superiority of the complex model over the simple model. Von Winterfeldt and Fischer (1975) state that "additive models usually provide excellent approximations to the subject's judgment." Huber (1974) suggests that "simple additive models perform equally well or better than complex model types," and Yntema and Torgerson (1961) report the additive model is appropriate and adequate for many practical situations. Members of decision making groups find the rationale of the simple model easier to understand and less expensive to utilize than more complex approaches.

Simple Expected Utility models contain a number of assumptions (Einhorn and McCoach, 1977): 1) utility functions for the attributes are linear; 2) the total utility of an outcome is an additive function of the utilities of the attributes making up that outcome; 3) levels of attributes are monotonic, i.e., more is always better than less or less is always better than more; 4) attributes are valued independently of one another; and 5) important attributes have not been left out. Keeney and Raiffa (1976) state that complex multiplicative models can be used successfully when violations of the above assumptions make linear models inappropriate. Goldberg (1968) notes that complex models are more sensitive to changes in utilities of individual attributes than are additive models.

Since no model has been found to be clearly more effective than another, Johnson and Huber (1977) suggest that the model most acceptable to the decision maker should be selected. A simple Expected Utility model was considered appropriate for this study of clinical management decisions. Applications of this model may be particularly acceptable to physicians for the following reasons:

1. The focus is on outcomes. In choosing management strategies, physicians commonly agree that there should be a positive outcome for the patient from the selected treatment. The use of outcomes as criteria for assessing quality care is recommended by a number of clinicians. Laxdal et al. (1978) suggest using measures of preventable morbidity, mortality, costs, and patient satisfaction in the evaluation of quality care programs, and report that physicians are more motivated to participate in programs where procedures are clearly related to measures of improvement in patient care and health status. Williamson et al. (1975) state that the "systematic considerations of relationships between medical care and outcomes can provide a crucial means of cutting through the enormous number of irrelevant variables so often included in the assessment of quality and continuing medical education." Outcomes can be examined within the decision analysis framework because they represent either a benefit or risk to the patient, and both a probability and a utility can be assigned to them (Kassirer, 1976).

2. Simple Expected Utility model applications fulfill a clinical need. They combine the probabilistic character of the medical environment with current knowledge, incorporate a way of balancing the benefits and risks to the patient of certain procedures, and explicitly outline the expected utility of various alternatives when considering choices in diagnosis and therapy. Childs and Hunter (1972) and Donabedian (1976) report these are key considerations that characterize the practice of a good physician and are at the heart of technical quality.

3. Applications of this model can be readily applied and easily understood, and thus allow for the realistic constraints of the clinical environment. How well these constraints are reflected in the applications will determine its acceptability. According to Kunreuther and Schoemaker (in press), examples of realistic constraints are: a) limited time, b) different cognitive abilities and styles, c) limited training in decision analysis, and d) varying needs. Clinical application of a simple Expected Utility model and communication of the significance of its results require less explanation, less training, and less time than would complex model applications.

Multi-Attribute Utility Theory

Any decision problem is characterized by the availability of more than one alternative course of action. In complex decision situations such as those that arise in a clinical environment, an outcome resulting from a chosen clinical strategy rarely can be described completely in terms of a single attribute. Outcomes described by more than one attribute are referred to as multi-attributed outcomes. Simple Multi-Attribute Utility Theory (SMAUT) models are designed to assess the combined effect of the outcome attributes in terms of the avowed preferences of the decision makers. To assess management alternatives by the use of SMAUT, one proceeds through the following five steps (Johnson and Huber, 1977; Edwards, 1977):

Step 1. The decision problem is structured. This step involves identifying the key decision makers, the decision alternatives, and the significant components or attributes of the outcome of a particular choice of action.

Step 2. The decision maker's preferences (utilities) are assessed for each of the identified attributes. Treatments have relative values for physicians depending upon benefit or risk (cost) to the patient. Treatment benefits may be measured in the degree of alleviation of suffering or reduction in morbidity or mortality. Costs may be assessed in terms of increased risk of morbidity or mortality, discomfort induced or increased, or monetary expense (Kassirer, 1976). There are several methods for assessing preferences for attributes (Johnson and Huber, 1977; Edwards, 1977; Weinstein et al., 1980; and Torrance, 1972). This study used a modified "direct method" or category scaling assessment technique (Edwards, 1977). This method requires the decision maker to rank all aspects of an attribute to be assessed, from the least to the most preferred. These least and most preferred aspects are then anchored at the extremes of an arbitrary utility scale ranging from 0 to 100. The evaluator chooses a numerical value for all intermediate attribute levels, somewhere between 0 and 100. There is a utility scale for each attribute of a given outcome. The utility scale for any one attribute is not necessarily comparable to the utility scale of another attribute, even though each ranges from 0 to 100. For example, the lowest level for one attribute may be valued higher than an intermediate level for another attribute. To overcome this difficulty, relative importance weights for each attribute are elicited from the decision maker. A weight of 10 is arbitrarily assigned to the least important attribute, and the other attributes are assigned importance weights in relation to this point. Importance weights for each attribute are normalized by summing them and dividing each weight by the total sum. Preference (utility) for a particular attribute level is calculated by multiplying the values assigned to various attribute levels by the importance weight placed on each attribute. The high face validity of this technique and the ease with which it can be taught to interviewers and decision makers make it a favored method for determining preferences. However, the method does not require the decision maker to make choices between alternatives and, therefore, cannot assess preferences between different choices as effectively as some other methods. Also, the arbitrary units used in the scaling may be unfamiliar to decision makers. The units must be translated to a scale of 0 to 1 to fit the decision analysis framework.

The standard gamble technique (Weinstein et al., 1980) and the time trade-off method (Torrance, 1972) are two alternative traditional approaches to assessing preferences for attributes. These procedures are detailed in Appendix A.

Step 3. The uncertainties (probabilities) are assessed for each of the identified attributes. Whether objective or subjective probabilities are used depends largely on the specific problem. Research reports note that subjective probabilities are quite satisfactory for use in clinical decision making where hard data are not available. In an unpublished doctoral dissertation, Gustafson (1966) compared the use in Bayes' Theorem of actuarially-based likelihoods and subjective probabilities to predict the length of patients' hospital stay. The subjective probabilities were found to predict the correct length of stay more often than the actuarial likelihoods. Lodwick (1966) demonstrated that when probabilities from a public health statistics model (Beach, 1975) are adjusted by incorporating subjective probabilities based on the personal experience of clinicians, the correct diagnosis of bone tumor is increased by approximately 5%. Fryback (1974) found that when explicit subjective probabilities of diagnoses are considered, the overall cost of chosen diagnostic strategies was decreased. In his study, five radiologists evaluated 50 cases of renal lesion and were asked to choose between the diagnostic strategies of percutaneous renal needle aspiration and renal arteriography to determine if the lesions were cysts, malignant tumors, or normal variants. The radiologists' costs of diagnostic strategy decisions were reduced on the average by 21% when they were asked to consider decision analysis principles, compared to the costs of these decisions without analysis. An additional 7% decrease in costs, on the average, resulted when the physicians were further asked to explicate their subjective probabilities of the possible diagnoses. The attributes considered in this study were dollar costs of the procedure, time lost by the patient from normal activities, patient discomfort, and risk of complication for the patient.

The ability to use subjective probabilities is also advantageous in that the collection of these estimates usually requires less time and expense, compared to collecting actuarial data.

Step 4. The alternatives are evaluated. The numerical values obtained in steps 2 and 3 for levels of each attribute are aggregated across attributes using a suitable aggregate and weighting scale. This mathematical combination results in a predictor score called a subjective expected utility (SEU) for each alternative course of action. This score can help determine choice of management alternative.

Step 5. An alternative is chosen. The alternative with the highest SEU is considered the best balance between risks and benefits.

Studies of Practice Patterns Without Decision Aids

Evidence has accumulated that physicians' clinical choices can vary in ways relatively independent of the clinical problem or patient health status (Childs and Hunter, 1972; Eisenberg, 1979; Eisenberg and Nicklin, 1981). Practice patterns are reported to be affected by physician attributes and characteristics such as specialty, years of practice experience, age, education, professional work environment, and doctor-patient relationships. They are also affected by patient characteristics such as social class, income, ethnic background, sex, physical appearance, and family influence. Some reports on the effects of physician characteristics on practice patterns are summarized below. Since patient characteristics are held constant in this study, the literature is not reviewed.

Childs and Hunter (1972) report on the influence of specialties on x-ray use. In their study of 800 sets of patients 65 years of age and older, they found that general practitioners (n = 80), internists (n = 44), and "other specialists" (n = 29) differed significantly in their patterns of x-ray use. Internists and "other specialists" were found to use x-ray more often than general practitioners. When diagnostic category was controlled, general practitioners (n = 131) and internists (n = 203) were found to differ in their choice of chest x-ray procedure. General practitioners more frequently used single views of the chest, while internists, except in the case of hypertensive cardiovascular disease, favored the use of fluoroscopy. Pulmonary disease, arteriosclerotic heart disease, and hypertensive cardiovascular disease were the three diagnostic categories reviewed.

Smith and McWhinney (1975) compared nine physician subjects each from the Departments of Family Medicine and Internal Medicine on diagnostic procedures through interaction with a simulated patient who presented three clinical problems. Internists asked a significantly greater number of history questions, elicited more data on physical examination, and ordered significantly more laboratory investigations than did family medicine physicians. The groups did not differ significantly in the final diagnosis.

Three rheumatologists and two clinical pharmacologists were found to differ widely in the relative importance each specialty attached to 14 agreed-upon areas needed to assess the severity of degenerative joint disease (Hammond and Joyce, 1977). They were also differed by specialty in their use of the cues in assessing 30 sets of clinical data. Within specialties, individual differences also occurred that seem unrelated to the clinical problem. Specifically, with respect to the rheumatologists in the preceding study, considerable variation in judgment consistency was demonstrated--90%, 31%, and 74%, respectively. The clinical pharmacologists, in contrast, were similar in judgment consistency--86% to 88%.

Koch-Weser (1977) studied the independent evaluations of three clinical pharmacologists on 500 cases of suspected adverse drug reactions. Although the pharmacologists based their evaluations on the same information and definitions, they differed widely in their judgments on whether an adverse drug reaction (ADR) had actually occurred, the ADR had caused the symptoms for which the patient was admitted, the drug indicated by the doctor was actually responsible for the ADR, the ADR contributed to a patient death, or on the role of drug interaction.

Specialty experiences can affect physician perceptions. Detmer et al. (1978) studied the questionnaire responses of 38 surgeons working in specialty groups characterized by high (2.42%) and low (0.44%) mortality rates. The study reports that the subjects' assessments of an overall surgical service mortality rate were biased by their specialty group. Specifically, surgeons in the high mortality group estimated the overall rate to be double that estimated by those from the low mortality group. The specialty groups consisted of 17 general surgeons, seven cardiac and thoracic surgeons, two urologic surgeons, three neurosurgeons, four orthopedic surgeons, and five plastic surgeons.

A 1981 study examined the use of laboratory tests and roentgenograms by three groups of community physicians. Internists and family practitioners were found to order a similar number of laboratory tests and roentgenograms when case mix was controlled (Eisenberg and Nicklin, 1981). General practitioners did not differ significantly from these two specialties in x-ray use, but were found to differ from the other two specialties in that they ordered fewer laboratory tests ($p < 0.01$). These findings were based on 55,420 outpatient visits to 336 physicians. The study controlled two other variables found to significantly affect x-ray and laboratory utilization—years of practice and medical school from which the physician graduated. Differences by specialty still persisted at the $p < 0.05$ level.

Many studies have noted the effect of the number of years a physician has been in practice. Stolley et al. (1972) surveyed 29 general practitioners, five osteopaths, and three internists about their prescribing habits relative to common complaints and illnesses, e.g., nausea, insomnia, uncomplicated common cold, and arthritis. These primary care physicians were also interviewed regarding their views on the use and contraindications for use of five specific drugs: Ritalin, Equagesic, Chloromycetin, Vitamin B12, and oral contraceptives. Evaluation by a panel of 33 expert judges of the quality of the physicians' behavior represented the study's dependent variable. The physicians were ranked from least to greatest appropriateness of prescribing. Physician characteristics, acting as independent variables, were tested as predictors of the physician's relative location in that ranking. The study found that "the younger, more recently trained physician with fewer years in practice, who goes on to take additional special courses or post-graduate training, is likely to be a more appropriate prescriber relative to his peers."

Two groups of physicians, one group graduating from medical school before 1946 and the other graduating from medical school in 1946 or later, were compared in the Childs and Hunter (1972) report. The later graduates referred their patients to radiologists significantly more frequently than did the physicians who graduated before 1946.

The 1981 Eisenberg and Nicklin study stated that:

> The number of years since medical school graduation showed a statistically significant inverse relationship with laboratory and radiology use. Laboratory tests per visit ranged from 0.26 for physicians who had graduated within the past 9 years to 0.02 for physicians who were 50-59 years out of medical school. Multiple analysis of variance showed the difference among groups to be significant ($p = 0.003$). Similarly, roentgenograms ranged from a mean of 0.066 procedures per patient

> visit for the cohort of most recent graduates, to 0.006 for those who graduated more than 50 years ago. Multiple analysis of variance showed the differences among the groups to be significant ($p = 0.008$). When case mix was controlled, similar results were found."

When the study controlled for confounding variables, physician specialty and medical school from which the physician graduated, the statistical significance remained at least at the $p < 0.05$ level.

Other factors are also reported to affect practice patterns. Childs and Hunter (1972) compared the diagnostic use of x-ray procedures by non-radiologists who provided direct patient x-ray services (group 1, $n = 153$) with non-radiologists who did not provide these services (group 2, $n = 610$). Group 1 x-rayed twice as many patients as group 2. In addition, the rate of use of x-ray procedures was 65% higher for group 1 than group 2 ($p < 0.001$). The authors suggest that two factors, ready access to x-ray equipment and economic interest, may account for these differences.

Eisenberg and Nicklin (1981) found that physicians who attended public medical schools ordered significantly more laboratory tests per visit than those who attended private schools, and foreign-trained physicians ordered more radiologic procedures than graduates of American medical schools. These findings were statistically significant after case mix, years since graduation, and specialty variables were controlled. Location of practice (urban, population greater than 100,000; rural, population less than 10,000; and intermediate areas), group or non-group practice, and medical or osteopathic education variables were found not to be significantly related to the use of diagnostic procedures.

The research cited above provides some evidence that factors other than the clinical problem and the health status of the patient may influence clinical decisions. The decision analysis technique enables the decision maker to focus upon factors that are directly relevant to the problem under consideration.

Research findings on the application of decision analysis techniques are summarized below.

Studies of Decision Making With the Aid of Decision Analysis

Multi-Attribute Utility Theory (MAUT) Model Applications

The Multi-Attribute Utility Theory (MAUT) model has been viewed from both a descriptive and prescriptive perspective. Some writers who report on the use of this model to describe and predict decision choices are Bierman et al. (1965), Einhorn and McCoach (1977), and Gardner and Edwards (1975). Others suggest that the model can only describe and predict in simple problem situations (Wendt and Vlek, 1975; Marschak, 1964). In complex situations, behavior resulting from the use of the model may not always correlate with behavior based on intuitive judgment. In fact, this difference may illustrate the power of the technique to improve on intuitive decision making (Howard, 1980).

Non-Clinical Applications

Klahr (1969) used a simple MAUT model with higher order functions to study the relation between intuitive and model based evaluations of student applications for college admission. The subjects, four males and one female, were members of the admissions staff at an undergraduate college of engineering and science. Each subject rated the student applications intuitively, and then a two-stage rating approach was used to collect responses in terms of the MAUT model. Evaluations made with the aid of the model were found to be highly related to the intuitive ratings of prospective students made by college admission officers (r = .94).

Hoepfl and Huber (1970) asked 11 graduate students and six professors in engineering to use a qualitative description of hypothetical faculty members to rate them on a 0-100 scale. Responses were then collected using a multi-attribute utility two-stage rating model composed of six attributes. The researchers found that the evaluations obtained using the model were accurate predictors of the actual evaluator ratings. Median correlations between model-based and intuitive judgments ranged from .87 to .98, with correlation declining as the number of attributes increased.

Neter and Williams (1971) applied the expected utility model using the standard gamble approach to a property insurance problem. Specifically, subjects were asked to decide how much insurance to buy. The intuitive preferences of 26 subjects (insurance agents, risk managers, and non-sales employees of an insurance firm) were compared to their choices determined by responses to the utility model. Only 3.8% of the subjects made the same choice with and without the aid of the model.

Gardner and Edwards (1975) studied the intuitive and multi-attribute utility responses of 14 individuals (two groups) involved in coastal zone planning who agreed to participate in the evaluation of 15 hypothetical realistic permit requests for development. The subjects made intuitive evaluations and also evaluations based on a multi-attribute utility theory (MAUT) procedure. The evaluations based on the model correlated with intuitive evaluation for both groups, .94 and .66, respectively. Following the initial comparison between intuitive evaluations and evaluations based on the model, each participant was provided feedback on the mean intuitive rating of his/her group, importance weights, etc. Values were discussed and clarified within groups, and a second intuitive evaluation took place. Group product-moment correlations between this second intuitive evaluation and the evaluation based on the model for both groups were .92 and .87, respectively.

Einhorn and McCoach (1977) utilized the simple multi-attribute utility procedure for evaluation of player performance in the National Basketball Association. Eight attributes of player performance were used as input to the model. The members of the actual first and second all-star teams for two seasons (1973-74 and

1974-75) were compared to those members who would make the teams if the model were used for member selection. The model predicted 13 of the 20 actual all-star team members.

Clinical Applications

A 1969 study by Stimson found that 11 members of a public health agency chose among federal grant applications as if they were maximizing expected utility. A model using attributes that represented the agency's grant application selection goals was designed to assist in the allocation. The application rankings derived from the use of the model were compared to the intuitive rankings. The model ranked 19 of the 22 first and second intuitive choices as first, second, or third in effectiveness.

Gorry et al. (1973) report on the use of decision analysis principles to design a computer program. The computer program was designed to select the most appropriate treatment for renal disease using 18 hypothetical clinical problems in which there were varying degrees of uncertainty as to the true diagnosis. "Two nephrologists agreed on the appropriate test or treatment for each of these cases, and their decisions served as a standard for evaluating the responses of the program. In 14 of the 16 cases, the decision of the program and physician agreed, and in the remaining four cases, the first choice made by the program was considered by the physicians to be a reasonable one."

Gustafson and Holloway (1975) applied a multi-attribute utility model to develop a burn severity index using responses from four physicians. Patient descriptions rank-ordered according to burn severity based on a model were highly correlated with the rank-ordered descriptions of 15 hypothetical patients rated on burn severity by the same four physicians (Spearman r = 0.89). The model's rank-ordered patient descriptions were also highly correlated with the ratings on the same hypothetical cases by eight physicians at two other burn centers. Spearman rank correlations with ratings collected at these two burn centers were 0.74 and 0.89, respectively.

Effect of Selected Physician Characteristics on Model-Derived Decisions

One general characteristic reported in using decision analysis models is that experts, because they focus on similar objectives and similar information, will show greater inter- and intra-judge agreement in their model-derived choices than in their intuitive choices (Einhorn, 1972; Aschenbrenner and Kasubek, 1978). Aschenbrenner and Kasubek (1978) found that multi-attribute evaluations of five experienced physicians using two preference assessment methods were in high agreement in their estimates of the overall dangerousness of seven cortizone drugs, while intuitive evaluations from these same decision makers were quite divergent. The two preference methods applied were a two-step rating procedure and an indifference curve method. Only 4.8% and 13.1% of the variance in judgments in the two-step and the indifference curve methods, respectively, were attributed to divergent evaluations of the relative danger of the drugs. 56.5% of the variance in intuitive judgments was attributed to divergent evaluations by the physicians.

While studies report higher inter-judge agreement on decisions derived from decision analysis than from intuitive judgment, research findings are mixed on whether or not these decisions are biased by particular decision making groups, professional experience, or other selected decision maker characteristics. In Giaugue's 1972 study, a multi-attribute utility model was applied to programs for the prevention and treatment of streptococcal sore throat and rheumatic fever. Using client explicated probabilities and utilities, Giaugue compared model-based choices on effective diagnosis and treatment strategies from five patients, two physicians, three nurse practitioners, and three public health officials. In this case, subjects' choices were not affected by their group definition.

Pauker (1976) studied decisions on coronary surgery through the use of a multi-attribute prognostic model. His study reported that decisions about therapy will be affected by differences in surgeons' past experiences with coronary artery disease management and by the differences in patient utilities. Coronary surgery was reported to be the generally preferred therapy for patients with disabling angina; it was rarely the preferred management choice for asymptomatic patients.

Krischer (1976) studied cleft lip and palate teams, 89 individuals from 17 facilities representing 13 specialties, and found that they made model-based therapy choices that were consistent with the favored regimen at their home facility.

Effect of Selected Decision Maker Characteristics on Parameter Estimations

Probabilities

Probabilities are one attribute parameter considered in the decision analysis model. Subjective probabilities are a decision maker's degree of belief about the occurrence of an outcome. Ginsberg and Offensend (1968) report in their study that there was considerable variation in the subjective prior probabilities elicited from two primary physicians on possible disease in a boy presenting with back pain and three collapsed vertebrae. The four diagnoses that were considered, with the physicians' probabilities, were: 1) infection, 0.15 and 0.45; 2) bone cancer, 0.25 and 0.45; 3) histiocytosis, 0.50 and 0.05; and 4) a rheumatoid nodule, 0.10 and 0.05. Physician characteristics that might account for this difference were not explored.

Six radiologists of differing experience and expertise varied in their probability estimates that a renal lesion was a cyst, a malignant tumor, or a normal variant; bias was present for all subjects. The biases in probability estimates were not related to professional experience (Fryback, 1974).

In Krischer's 1976 cleft lip and palate study, clinicians from a facility using presurgical orthopedics and from a facility using purely surgical intervention were asked to estimate the probabilities of cosmetic effects, speech clarity, and hearing improvement for these two management alternatives. The study reports that the physicians within each facility gave similar estimates of the subjective probabilities of each of the attributes. However, there were substantial differences on the physicians' estimates between the two facilities.

Importance Weights

In the decision analysis model, importance weight is a second attribute parameter. A weight of 10 is arbitrarily assigned to the least important attribute, and the other attributes are assigned importance weights in relation to this point. There is some evidence that the use of importance weights improves the accuracy of the prediction of the model. In their study of 15 experienced and 15 non-experienced persons seeking professional employment, Huber et al. (1971) found that a two-stage rating model using importance weights was superior to the unweighted rating models in predicting job choices in 27 of 30 subjects. Similar findings were reported by Vroom (1966) who studied 49 students enrolled for the masters of science degree in a graduate school of industrial administration. There was a marked positive relation between ratings of the attractiveness of organizations and scores on an instrumentality goal index that were obtained by combining data on the relative importance of different goals to the subjects and their perceived likelihood of obtaining these goals in a particular organization. Seventy-six percent of the subjects chose the employment with the highest instrumentality-goal score.

In contrast, other researchers (Einhorn and Hogarth, 1975; Dawes and Corrigan, 1974) have had satisfactory results from models which assumed attributes to be of equal weight. Mikes and Hulin (1968) report that importance weights had little value in predicting job turnover during an 11-month period. "Unweighted sums of the satisfaction scores alone proved to be as highly related to termination decisions as the sums of the importance-weighted satisfaction scores." The study conclusions were based on questionnaire responses of 660 male and female corporate office personnel.

Literature indicates that when several people are involved in making a decision, they will tend to agree on the attributes that are relevant but will vary on the importance weight each assigns to these attributes (Edwards, 1977; Hogarth, 1980). The different weights attached to different attributes may be the primary source of difference between evaluators. In Fryback's 1974 study of the diagnosing of a renal lesion, six radiologists, one neurologist, and two non-medical subjects assigned differing importance weights to the attributes of patient discomfort, medical risks, and amount of time lost from normal activities. Subjects all agreed in assigning no importance to the "dollar cost" of each procedure.

Krischer (1976) found that significantly more fathers than mothers or plastic surgeons working on cleft lip and palate patients weighted "speech intelligibility" higher than "cosmetic appearance." The authors also found that clinicians at a center using presurgical orthopedics put heavy weight on "cosmetic appearance," while clinicians from a center that favored purely surgical intervention weighted "intelligibility of speech" highest. For all subjects, the importance weights for speech clarity and cosmetic appearance were found to be very similar, as were the importance weights for monetary expense and hearing improvement.

Aschenbrenner and Kasubek (1978) report that importance weights provided the main source of physician evaluative information in assessing the overall dangerousness of alternative drugs.

Values

In decision analysis, values, the third attribute parameter, are assigned by the decision maker to various levels of each attribute, with 0 indicating the least valued level and 1 indicating the most valued. Intermediate levels are assigned values by one of the three preference assessment methods previously noted. In Ginsberg and Offensend's 1968 study of the boy presenting with back pain and collapsed vertebrae, the outcome utilities of the two primary physicians were diverse for four outcomes and similar for six. Specifically, "values of 25 and 94 for u(cure/six months); -150 and 40 for u(paralysis); and 80 and 100 for u(cure/1 month) were diverse. The two physicians were similar in their utilities for u(kyphosis), u(cure/now), u(cure/1 week), u(cure/histiocytosis), u(cure/cancer), and u(cure/death)."

In Giaugue's 1972 study using patient, physician, nurse practitioner, and public health official subjects, variations in utilities within groups exceeded any systematic differences between groups.

Fryback (1974) asked subjects to consider the potential consequences of four actions, needle aspiration of a renal cyst; arteriogram on a renal tumor or a normal kidney; both procedures on a renal tumor or normal kidney; and both procedures on a renal cyst, and assess each action on the four attributes of patient discomfort, complication risks, monetary cost, and time lost. Fryback discovered that wide variation existed in the average values assigned by his subjects. The values were assessed using a category scaling method. The intermediate values elicited were checked by a lottery approach.

In his study of management strategies for cleft lip and palate, Krischer (1976) used a form of the standard gamble or lottery assessment method to find that differences in utility functions for both monetary gain and monetary expense were statistically insignificant across groups of clinical specialties and family members. In addition, systematic differences in attitudes toward risks between a facility's clinical staff and members of its patients' families were not statistically significant. Utility functions for speech (percent of words from a sample of speech found to be intelligible) yielded no statistical differences when clinical specialty, family membership, or affiliation with a particular treatment center were compared as groups. However, when family members and selected clinical specialists considered the validity of the attribute independence assumption, some differences were significant.

> Significantly fewer fathers than mothers (probability $p < 0.05$), pediatricians ($p < 0.025$), and plastic surgeons ($p < 0.01$) felt that degree of speech intelligibility (paired with cost) was preferentially independent of the other attributes. Also, the proportion of speech pathologists from whom the preferential independence of speech and cost was valid was significantly less than either pediatricians or plastic surgeons ($p < 0.05$). Statistically significant differences were also found among the respondents when considering the pair-wise preferential independence of hearing and cost. Significantly more fathers than mothers ($p < 0.05$), pediatricians ($p < 0.025$), and otolaryngologists ($p < 0.05$) felt that hearing paired with cost should not be preferentially independent of the other attributes.

Krischer (1976) reported that for fathers and some speech pathologists, the utility of trading off money for speech and hearing was dependent on the levels of the other attributes.

Pliskin et al. (1980) administered a questionnaire concerning the willingness to sacrifice longevity for relief from angina pain to ten subjects age 27-55. The subjects were members of the Faculty Seminar on the Analysis of Health and Medical Practices at the Harvard School of Public Health. Three respondents were physicians; seven were economists and statisticians. All respondents were reported to be familiar with the medical and behavioral significance of anginal pain. The study's findings indicate that the respondents varied considerably in the way they valued pain and disability versus longevity. The respondents did not vary systematically with their age, number of children, or profession. The authors note the small sample size.

Action strategies derived from the decision analysis model vary according to other factors. Several studies report the effects of incorporating patient preferences, patient and family attitudes, or patient conditions or situations on the ultimate choice of management strategy. Henschke and Flehinger (1967) applied decision theory to the management problem of a patient with squamous cell carcinoma in the anterior two-thirds of the tongue. The usual management alternatives are prophylactic neck dissection or no prophylactic neck dissection. By adding a more precise description of the patient's condition to the decision equation, the authors demonstrated that the best results may be expected when prophylactic neck dissection is carried out only in patients with primary cancer larger than two centimeters.

Bunch and Andrew (1971) demonstrated through decision analysis that clinical management strategy choices may differ according to the patient's occupation, i.e., the choice of management strategies for a young healthy male adult presenting with a mid-shaft fracture of the femur may depend upon whether the patient's occupation is sedentary or heavy physical work.

In Pauker's 1976 study of coronary artery disease, the decision alternatives of coronary bypass surgery or medical therapy were posed. Patient preferences between length of life and quality of life were assessed, and the ultimate decision choice was highly affected by these preferences and attitudes.

Similarly, in a study addressing the coronary artery bypass surgery decision, Weinstein et al. (1977) found that a decision analysis that incorporated consideration of patient life style would recommend surgery for an active patient and medical management for more sedentary patients.

In a 1978 study, McNeil et al. (1978) demonstrated that incorporating a patient's attitude toward risk or longevity can influence the choice between surgery and radiotherapy in the treatment of lung cancer.

Again, in 1979, McNeil and Pauker showed the effect upon management decision choices of incorporating patients' attitudes on length of life. The authors interviewed 14 patients under treatment for bronchogenic carcinoma and incorporated these patients' attitudes into the evaluation of two diagnostic alternatives. Their study results indicate that exhaustive preoperative diagnostic testing to search for metastases is the decision of choice for some patients who are highly averse to the risk of operative death. The second alternative, surgery only if a preoperative metastic workup was negative, is based upon the five-year survival or life expectancy of patients, but does not incorporate or consider patients' attitudes.

Obesity

Multi-attribute theory can be directly applied to clinical management choices with respect to the condition of obesity, the health problem which this study addresses. A summary of the literature review on obesity is reported below.

<u>General Morbidity</u> (Effect of Condition on Physical, Social, and Emotional Health)

Obesity is a prevalent health problem in America. The National Center for Health Statistics data for 1971-74 (Abraham and Johnson, 1979) report that 14% (8,041) of men age 20-74 and 23.8% (15,268) of women age 20-74 are 20% or more above desirable weight. Desirable weight was defined as the mean weights for men and women aged 20-29 years.

Bray (1976) identified important health consequences associated with excess weight:

1. Significantly increased mortality when body weight is 25-30% or more above "desirable weight."

2. Obesity effects the functions of various organ systems, i.e.:

   "a. Obesity increases the work of the heart--cardiac output; stroke volume and blood volume are all increased.

   b. The pulmonary function is impaired--with gross obesity, severe alveolar hyperventilation may develop.

c. The endocrine function is modified--Beta cells in the islets of Langerhan are enlarged. Glucose tolerance may be impaired.

d. The pancreatic function may become impaired, causing hyperinsulinemia.

e. Abnormal menstrual cycles may occur, or even amenorrhea.

f. The release of growth hormone from the pituitary gland may decrease.

g. The secretion of adrenal steroids from the adrenal cortex and the secretion of 17-hydroxycorticosteroids in the urine are increased.

h. The concentration of triiodothyronine shows a small but significant positive correlation with overweight (thyroid function)."

3. Social and economic consequences result from obesity. Overweight or obesity tends to evoke negative affect and rejection. Labelled as socially undesirable, obese individuals often are considered social "misfits" and consequently may experience difficulty in obtaining employment or getting admitted to college (Bray, 1976).

Van Itallie (1979) reiterates Bray's concerns and notes that obesity is thought to contribute to health disorders in the following body organs or systems: heart, vascular, respiratory, hepatobiliary, kidney, skin, joints, muscles, and connective tissue systems. In addition, obesity is believed to increase the risk of endometrial and breast cancer; interfere with reproductive, sexual, hormonal, and metabolic functions; impair psychosocial functions; increase accident proneness; and interfere with the diagnosis and therapy for other disorders. Of all age groups, the obese middle-age population is most at risk for morbidity and mortality. Obesity is more prevalent in females than in males.

Results of studies which have examined the effects of obesity on a person's health status are reported. Van Itallie (1979) distinguished between the adverse effects on health of severe obesity (130-200% of averaged desirable weight), massive or morbid obesity (200% or more of desirable weight), and mild to moderate obesity (110-130% of desirable weight). Several individual studies that examined the effects of various degrees of overweight on mortality and morbidity are specifically reviewed.

Rimm et al. (1975) studied the relationship between obesity and disease within 72,532 women enrolled in TOPS (Take Off Pounds Sensibly) and living in the United States and Canada. The authors compared the frequency of disease in persons at five different obesity levels. The degree of obesity was measured in the ratio of weight to height. The scale for this ratio was divided into five intervals. The five-point scale was constructed to include 20% of the population at each obesity level. Relative risks were calculated for each disease condition by calculating a ratio which reflects the risk of having a history of the disease for the group with the highest weight as compared with the risk for the group with the lowest weight. The researchers assumed that risks reflected the strength of the relation between obesity and each reported disease. Results suggest that severe obesity in women between 30 and 49 years of age increases the risk of diabetes 4.5 times, high blood pressure 3.3 times, gallbladder disease 2.7 times, and gout 2.56 times.

Kral et al. (1977) and Weisinger et al. (1974) also report morbidity and mortality to be related to, or to be a complication of severe or massive obesity. Kral et al. (1977) studied the relation between obesity and hepatic steatosis in 17 grossly obese subjects (4 men, 13 women). The subjects had hyperinsulinemia, increased hepatic synthesis of lipids, increased hepatic content of lipids, and increased serum levels of free fatty acids. Four massively obese patients were found by Weisinger et al. (1974) to have positive health consequences, i.e., decreased proteinuria and coincidal reduced right arterial pressure and blood volume, during dietary weight loss. Van Itallie (1979) also cites research reporting that even the moderately obese patient is predisposed to peripheral venous stasis with varicose veins, hemorrhoids, and increased risk of thromboembolism.

Quereshi (1972) studied 180 obese women who failed to remediate their obesity. They found that the subjects perceived themselves as generally unhappy, nervous, tense, dissatisfied, lonely, and rejected. Effective treatment has been found to improve social and personality outlooks of patients. Solow et al. (1974) report that two years following intestinal bypass surgery, 29 massively obese male and female patients who lost weight reported an improvement in mood, self-esteem, and interpersonal and vocational effectiveness. They also reported a trend toward decreased anxiety and depression. These changes were directly proportional to magnitude of weight reduction. Collingwood and Willett (1971) also report increased self-esteem after weight loss. Specifically, the authors report the effects of physical training on five male teenagers enrolled in a three-week obesity program. The subjects experienced a significant weight decrease ($p = .005$) and reported a significant increase in self-concept and self-acceptance ($p = .05$).

The patient's age and age of onset of obesity are significant when discussing the effect of obesity on patient morbidity and mortality. Studies report that obesity acquired between the ages of 20-40 has a much greater effect on the development of subsequent cardiovascular disease than obesity which occurs after the age of 40. They also suggest that the obesity of middle-age is more detrimental to health and survival than obesity in old age (Van Itallie, 1979). Rabkin et al. (1977), in a longitudinal study, kept records on 3,893 men with a mean age of 30.8 and reported that body mass index was found to be a significant predictor of the 390 cases of ischemic heart disease that occurred over a 26-year period. The study concluded that obesity occurring between the ages of 20-40 is a definite risk factor.

Drenick et al. (1980) calculated the excess mortality in 200 obese men classified as morbidly obese (200% or more of desirable weight). They report that in those men 25-34 years old and 35-44 years old, the mortality ratios were almost 12 and 6 times the mortality ratios of males in those age categories in the general population.

Treatment or Management Strategies

General

Practitioners have long been faced with the problem of how to treat obese patients so that weight reduction occurs and is maintained over a long period of time. Maddox et al. (1966, 1968) report pessimism and indifference on the part of both the physician and the patient about the possibility of successful weight reduction, even when it appears to be warranted medically. Many factors, beyond medical health, appear to play a part in decisions related to management strategies or lack of them. Proposals for management are also related to physical and social-psychological characteristics of the patient. The probability of a specific management strategy being recommended increases as patient excess weight increases. An important factor is the perception of "obesity" by the public as a whole. In general, the public views obesity as evidence of self-indulgence and weakness. In addition, the condition is seen as physically unsightly. Physicians, as members of the public, share these attitudes. These negative feelings are also reinforced by their professional experiences dealing with the consequences of obesity. For example, surgeons are disturbed by the interference of overweight in surgical procedures, and internists see the harm or benefits to diabetics and hypertensives depending upon whether or not obesity is controlled (Keys, 1955). Eisenberg (1979) notes that medical students have judged overweight patients to be less likeable and to have a poorer prognosis. In spite of the negative attitudes of many of the medical profession toward obesity and the frustrations associated with its management, Bray (1972) reports that one of the primary responsibilities of a physician who undertakes to treat obesity is to motivate the patient and to provide continuing sympathetic understanding for accompanying emotional and medical ills.

Management Modes

There are several modes of management a physician may choose or recommend in obesity management.

Medical, surgical, psychotherapeutic, and dietary methods of weight reduction have been used with varying degrees of success. Treatment results are frustrating, however, in that many weight loss procedures are no more effective now than they were 20 years ago (Leon, 1976). Behavioral methods, therapist-provided rewards, contingency contracting, self-reward and self-monitoring, aversive procedure, covert sensitization, and coverant control* have been employed. Other psychological methods such as group support therapy, individual psychotherapy, and hypnosis have also been used. Abramson (1977) reports that self-control treatments are the most consistently successful. The typical components of this treatment include self-monitoring, self-reward and punishment,

---

*Contingency Contracting - "A procedure which involves an agreement between therapist and client on a reward-penalty system that is contingent on the client's weight changes. The client generally deposits some money or valuables with the therapist which are earned back or permanently lost." (Leon, 1976)

Self-Reward and Self-Monitoring - "Strategies that provide immediate incentives for improvement by allowing the individual to present himself/herself with a specified consequence immediately after the occurrence of a target behavior." (Leon, 1976)

Aversive Procedure - "Electric shock presented when approaching a craved-for food or favorite foods paired with highly noxious odors." (Leon, 1976)

Covert Sensitization - "Client is placed in a state of relaxation and develops an avoidance respondence through imagining the undesirable stimulus (eating) paired with an extremely aversive stimulus." (Leon, 1976)

Coverant Control - "Involves the systematic use of thoughts, images, and reflections to modify eating behavior." (Leon, 1976) and stimulus control techniques intended to reduce the number of environmental stimuli that trigger eating.

One mode of treating obesity is referral to an endocrinologist for diagnostic or management purposes. Endocrine alterations, such as Cushings syndrome, insulinoma, the empty sella syndrome, hypogonadism, hypothyroidism, hypoparathyroidism, and pseudo-hypoparathyroidism, are rare causative factors for obesity (Bray, 1976). However, according to a recent chart audit study that specifically examined primary care physicians' decisions to refer or not to refer obese patients to an endocrinologist, obese partients even with low probabilities of endocrine disease are referred to endocrinologists (Ravitch et al., 1982). This study's findings were based on the analysis of 83 medical records of female and male patients referred to an endocrine clinic and 300 obese patients randomly drawn from the general ambulatory clinic record files.

In sum, obesity is a serious and widespread medical problem, and numerous medical, psychotherapeutic, and dietary methods of weight reduction have been devised to treat this condition. These methods have been used with varying degrees of success in the attempt to reduce weight and maintain that loss over time. Obesity is more prevalent in females; the middle-aged are most at risk for morbidity and mortality.

Referral to an endocrinologist is one possible treatment strategy for primary care physicians in the management of obese middle-aged female patients. It is assumed that weight reduction and patient satisfaction with medical management are two goals that guide a physician's choice of management strategy.

Decision analysis can be applied to this particular medical condition. The management alternatives of referral and non-referral can be examined in reference to the goals of weight reduction and patient satisfaction with medical management. These attributes have been defined as adequate to incorporate the important aspects of the problem. They are considered operational as they are meaningful to the physician and to the patients. The two attributes can be decomposed into levels so that their valuing can be handled intelligibly and are limited in number, as recommended by Keeney and Raiffa (1976) and Edwards (1977). Following the recommendations of previously published research, a simple linear multi-dimensional utility procedure will be used to analyze the obesity management problem. The specific details of the study methodology and procedure are described in Chapter III.

CHAPTER III

STUDY METHODOLOGY AND PROCEDURE

In this study of clinical decision making, physicians responded to a series of hypothetical standardized vignette cases and a semi-structured interview on the question of referring obese patients to an endocrinologist.

Selection of Subjects

Physicians in the specialties of General Internal Medicine, Family Practice, and Gynecology are self-defined as providers of primary care (Rothert et al., 1982). It was assumed that primary care physicians see a significant number of obese patients in their practices. The primary criterion used for subject selection, therefore, was that the subjects be within these specialty groups. Subjects were recruited primarily from the greater Lansing area. Selection continued until 15 volunteers, 12 males and 3 females, had been recruited from each specialty. These subjects represented approximately 50% of the Gynecologists, 100% of the General Internists, and 17% of the Family Practitioners in the greater Lansing area.

Study participants were recruited in several ways. An endocrinologist and a gynecologist associated with decision analysis research identified colleagues, and these colleagues were asked at the termination of their interview if they knew of primary care physicians who might be willing to participate. The telephone directory also was a source of names of physicians to contact. Finally, the gynecologist on the research team presented an outline of this study at an Obstetrics/Gynecology meeting in Grand Rapids and called for volunteers. The Chairman of Family Practice did the same with physicians in the Family Practice unit at Michigan State University.

Subjects were contacted by telephone or in person to set up appointments for interviews. A cover letter confirming the time and date of the interview was mailed or delivered in person to each subject, along with the vignette booklet.

Fifty-five primary care physicians were contacted in order to recruit 45 subjects. Ten physicians were unable to participate because of inability to schedule interviews within the data collection time frame, illness, or reasons unstated. The volunteer subjects received $100 for their participation. Anonymity of the subjects was guaranteed. Subjects were given the opportunity to request a summary of the results if they so desired. Characteristics of the subjects are summarized in Table 1.

Table 1

Characteristics of Research Subjects

		FP	IM	OB/GYN	Totals
M.D.	M	7	11	10	28
	F	1	3	3	7
D.O.	M	5	1	2	8
	F	2	0	0	2
		15	15	15	45

Subjects had a mean age of 39.5 and mean number of years in practice of 12.6 (8.7). The average number of years in practice for Family Practitioners and General Internists was similar ($\bar{X}$ 10.4, S.D. 8.7; and $\bar{X}$ 10.8, S.D. 8.1, respectively). For Gynecologists it was considerably greater ($\bar{X}$ 16.6, S.D. 8.5). The average number of years in practice was higher for male than for female physicians. The males averaged 14.0 (S.D. 9.1) years in practice; the females averaged 6.9 (S.D. 3.8) years. Thirty-three subjects represented community-based practices and 12 were academically based, as defined by the setting in which they see patients.

Rationale for Sample Size

Several factors were considered when determining the sample size: the accessibility and availability of the study population, the cost of gathering data, and the number of subjects necessary to detect significant differences. According to Cohen (1977), with 45 subjects, 15 in each specialty group, a difference of one standard deviation between two groups can be detected with .74 power (alpha = .05) and a difference of one standard deviation among three groups can be detected with a .78 power.

Controls for Contamination

As subjects were not selected randomly from the defined primary care physician population, steps were taken to control for contaminating factors when examining differences by specialties. Previously cited research (Stolley et al., 1972; Eisenberg and Nicklin, 1981) indicates that the amount of experience in practice influences management decisions. As it was not possible to match the specialty group subjects on this variable in the selection process, the statistical technique of analysis of covariance was used to control for the initial differences in the sample. The groups were comparable with respect to gender, containing 12 males and three females from each represented specialty.

Referral Behavior as Described by the Application of the Multi-Attribute Utility Model

The study used a simple linear multi-attribute utility model, incorporating five steps of Hogarth's (1980) seven-step framework. The relevant steps for this study are:

1. Structure the problem,
2. Assess preferences (utilities) for each attribute,
3. Assess uncertainties (probabilities) for each attribute or attribute level,
4. Evaluate alternatives, and
5. Choose an alternative.

These steps are detailed below.

Step 1: Structure the Problem

Overview: Presuming that an endocrinologist is available who routinely manages obese patients, will the primary care physician generally choose to refer the obese female patient (50% and 100% overweight and presenting no physician cues associated with endocrine disease) to the endocrinologist, or will he/she choose to not refer the patient?

a. Key decision makers: primary care physicians who see and manage female obese patients (50% and 100% overweight) in their offices.

b. Decision alternatives:

    1. to refer to an endocrinologist, and
    2. not to refer to an endocrinologist.

c. Attributes that affected choice of decision alternatives:

    1. Weight Reduction

        Three levels were conceptualized in the model:

        Little or no weight reduction (W_O)

        Weight reduction of approximately one-half the excess weight (W_M)

        Weight reduction to approximately ideal weight (W_I)

    2. Patient Satisfaction

        Two levels were conceptualized in the model:

        Patient satisfied (PS)

        Patient not satisfied (NS)

d. Key uncertainties for this model concern the probability of each level of weight reduction and patient satisfaction:

1. Weight reduction - little or none
2. Weight reduction - approximately one-half the excess weight
3. Weight reduction - approximately to ideal weight or 10-15% above
4. Patient satisfied
5. Patient not satisfied

The clinical decision tree for this situation is presented in Appendix B.

Step 2: Assess Subjects' Preferences (Utilities) for Attributes

A modified direct method for assessing preferences was selected because subjects could respond easily and interviewers require little training (Johnson and Huber, 1977). As preference (Ui) for a particular attribute has two components (importance weights, Wi; and value, Vi), a direct method two-stage assessment technique was used. First, importance weights were assessed for each of the model's attributes, i.e., patient satisfaction (PS) and weight reduction (WR). The values were then determined for each level within each of the two attributes. The procedures are detailed below.

Assessing Importance Weights (Wi)

Importance weights for each attribute were determined by first asking the physicians to rank attributes in order of importance and assign a weight of 10 to the least important. Then they were asked to evaluate how much more important they considered the second attribute by assigning it a number which represented the ratio of its importance to the anchor point 10. This assessment technique was tested in two pilot studies to ensure that it was appropriate to the situation and acceptable to the decision makers. When the second number had been assessed, the subjects' importance weights for each attribute were normalized by summing them and dividing each by their sum. For example, if the second attribute were judged to be twice as important as the first, it would be assigned a weight of 20, and the normalized importance weights would be .67 (20/30) and .33 (10/30).

Assigning Values (Vi)

The category scaling method was used to assess values of the levels for the attributes, weight reduction and patient satisfaction.

a. Values Estimated for Levels of Weight Reduction (V_1)

For each physician subject, the most preferred attribute level (W_I--patient reaches 10-15% of ideal weight) was arbitrarily anchored at a value of "1." The least preferred attribute level (W_O--patient loses little or no weight) was arbitrarily anchored at a value of "0." According to the principles of utility assessment, the value of the intermediate level, W_M (weight reduction of approximately one-half the excess weight), must fall between the values "1" and "0" and must be assessed by the subject. Excess weight is medically undesirable because it increases morbidity; therefore, each subject's estimated value of the intermediate level, $V(W_M)$, was derived from estimations of the increase in patient morbidity associated with stated weight levels in the vignettes. Twelve case vignettes described patients who were approximately 100% overweight; 12 described patients approximately 25-50% overweight. After each vignette, each subject was asked to respond on a seven-point scale to this question: "How does this patient's obesity increase her chances of morbidity?"

Each subject's mean response was calculated for the 12 100% overweight patients and for the 12 25-50% overweight patients. For each physician, the mean response for the 12 100% overweight patients was considered to be a measure of estimated disutility($\bar{d}$) for being or staying at 100% overweight $\bar{d}(W_O)$; the mean response for the 12 25-50% overweight patients was taken to be the disutility estimated for being 50% overweight $\bar{d}(W_M)$. The mean estimate of disutility for the 100% overweight group was then contrasted with the mean estimate of disutility for the 25-50% overweight group to determine the disutility of being at the intermediate weight level.

The disutility for $W_M = \bar{d}(W_M) = \dfrac{\bar{d}\,(W_{50}\ \text{o.w.})}{\bar{d}\,(W_{100}\ \text{o.w.})}$

The value for moderate weight loss (W_M) was defined as:

$$1 - \bar{d}(W_M) = \bar{v}(W_M) = 1 - \frac{\bar{d}\,(W_{50}\ \text{o.w.})}{\bar{d}\,(W_{100}\ \text{o.w.})}$$

For example, if a subject's mean response for the 12 100% overweight patients was 6.83 and the mean response for the 12 25-50% overweight patients was 4.75, the calculations for the derived $\bar{v}(W_M)$ would be as follows:

$$\bar{v}(W_M) = 1 - \frac{\bar{d}\,(W_{50})}{\bar{d}\,(W_{100})}$$

$$\bar{v}(W_M) = 1 - \frac{4.75}{6.83} = .30$$

Similar procedures were carried out to determine the $\bar{v}(W_M)$ for each subject for 50% overweight patients (n = 8) by comparing morbidity ratings of these patients to the mean morbidity ratings of the 25% overweight patients (n = 4).

The procedure described above for assessing values for the intermediate levels of the weight reduction attribute, $\bar{v}(W_M)$, was devised to avoid an overly long data collection procedure which might have decreased the quality of the data. In the opinion of the decision making research team, which included two primary care physicians, the values for $\bar{v}(W_M)$ for each subject could be adequately defined in terms of that physician's estimate of the increase in patient morbidity associated with stated weight levels in the vignettes.

b. Values Estimated for Levels of Patient Satisfaction (V_2)

This attribute was arbitrarily assigned two levels, patient satisfied (PS) and patient not satisfied (NS). The most preferred outcome, PS, was anchored at a value of "1," and the least preferred outcome, NS, was anchored at a value of "0."

Step 3: Assess Uncertainties

The subjective probability estimates, P_i, were determined by presenting referral and non-referral scenarios to the physicians during an interview. They were asked to imagine they had 100 (100%) overweight female patients and to estimate the percentage of these patients who will have lost no weight, about half the overweight, or reduced to ideal weight in two years time, if the patients were referred to an endocrinologist. The questions were repeated for 100 (50%) overweight patients, and both series were repeated for the management alternative of retaining the patients without referral. A similar series of questions were asked the subjects for both 50% and 100% overweight patients on probability estimates of patient satisfaction or dissatisfaction after two years, if referred to an endocrinologist or managed by the primary care physician.

Step 4: Evaluate Alternatives

The subjective expected utility for each alternative, referral and non-referral, were calculated for both the 50% and 100% overweight cohorts by the following mathematical procedure:

$$\underset{\substack{\text{subjective expected}\\ \text{utility for the jth}\\ \text{subject}}}{SEU_j} = P_1U_1 + P_2U_2 + \ldots + \underset{\substack{\uparrow\\ \text{probability}}}{P_i}U_i \longleftarrow \begin{array}{l}\text{preference:}\\ \quad\text{value } (V_i)\\ \quad\text{importance weights } (W_i)\end{array}$$

j = physician 1-45
i = attribute 1 or 2

Therefore, for this problem the formula became:

$$\mathrm{SEU}_j = \underset{\substack{\uparrow \\ \text{relates to weight} \\ \text{reduction attribute}}}{P_1W_1V_1} + \underset{\substack{\nwarrow \\ \text{relates to patient} \\ \text{satisfaction attribute}}}{P_2W_2V_2}$$

The formula was expanded further since each attribute had more than one level. Therefore, $P_1W_1V_1$, because of three levels of weight reduction, became:

$$P_{(WO)}W_{(WR)}V_{(WO)} + P_{(WM)}W_{(WR)}V_{(WM)} + P_{(WI)}W_{(WR)}V_{(WI)}$$

OR

$$W_{(WR)}\,(P_{(WO)}\cancelto{0}{V_{(WO)}} + P_{(WM)}V_{(WM)} + P_{(WI)}\cancelto{1}{V_{(WI)}}) = W_{(WR)}\,(P_{(WI)} + P_{(WM)}V_{(WM)})$$

$P_2W_2V_2$, because of two levels of patient satisfaction, became:

$$P_{(PS)}W_{(PS)}V_{(PS)} + P_{(NS)}W_{(NS)}V_{(NS)}$$

OR

$$W_{(PS)}\,(P_{(PS)}\cancelto{1}{V_{(PS)}} + P_{(NS)}\cancelto{0}{V_{(NS)}}) = W_{(PS)}\,(P_{(PS)})$$

SEU_j formula now became:

$$\mathrm{SEU}_j = W_{(WR)}\,(P_{(WI)} + P_{(WM)}V_{(WM)}) + W_{(PS)}\,(P_{(PS)})$$

Step 5: Choose an Alternative

It is expected that the decision maker will choose to refer or not to refer the patient depending upon which action scored the highest expected utility.

In this study, a net subjective expected utility for both 50% and 100% overweight patients was calculated for each subject to determine the degree to which referral or non-referral was preferred. This was done by subtracting the elicited expected utility for non-referral from the expected utility for referral for each subject. The net subjective expected utility (Net SEU) was represented symbolically as Net SEU = SEU(R) - SEU(NR).

Net SEU's may vary in numerical value from +1 to -1. A positive net SEU indicates referral has the higher expected utility for the decision maker; a negative net SEU indicates non-referral has the higher expected utility; a net SEU of zero implies that the decision maker values referral and non-referral equally.

Assumptions (Einhorn and McCoach, 1977) that are made when the multi-attribute utility model is applied are that the values assigned to the attributes are linear; that the total value of an outcome is an additive function of the values of the attributes making up that outcome; that the levels of the attributes are monotonic, i.e., more weight loss would always be better than less; that attributes are valued independently of one another, i.e., patient satisfaction with management is independent of weight loss; and that important attributes have not been left out.

Referral Behavior as Described by Responses to Case Vignettes

Physician responses to case vignettes were used to describe clinical referral behavior. The data were collected by questions that followed each case vignette asking the physician to estimate the probability that he/she would refer the patient depicted in the case to an endocrinologist. The subjects responded by choosing a position on a seven-point scale with the intervals defined as follows:

1. Virtually certain <u>not</u> to refer (0-5%)
2. Fairly certain <u>not</u> to refer (6-20%)
3. Probably would <u>not</u> refer (21-40%)
4. May or may not refer (41-60%)
5. Probably would refer (61-80%)
6. Fairly certain to refer (81-95%)
7. Virtually certain to refer (96-100%)

A case was identified as referred if the subject circled 5, 6, or 7 (a 60% or greater chance of referral).

Physician responses to case vignettes were used to describe clinical behavior rather than information from chart review or direct observation of physicians in their offices because a major purpose of this study was to examine variability in performance due to physician characteristics. With standardized cases, systematic differences can be attributed more confidently to physician characteristics since the confounding variable of patient differences has been controlled.

Instruments

Description

Two instruments were chosen for use in this study: 1) brief written structured case vignettes, and 2) a semi-structured interview. Both instruments were developed in the Fall of 1980. A research team participated in the intial development and revision of the instruments. Expertise within this group included primary care physicians, decision analysis experts, psychologists, measurement design experts, health professionals, and an endocrinologist. This investigator played an important part in the design and revision of the instruments.

The general structure of the case vignette booklet included a brief explanation of the study, instructions, a definition of obesity, and 32 randomly presented vignettes. Twenty-four of the cases were unique, six were replications, and two were practice cases. Twelve of the cases concerned patients approximately 100% overweight and 12 depicted patients approximately 50% overweight. The vignettes reflected the difficult management problems of obese female patients who do not present physical cues associated with endocrine disease. A copy of the explanation and directions in the vignette booklet and examples of the case vignettes are in Appendix C.

A semi-structured interview format was chosen to gather and check on responses to questions about probability and importance weight estimations since it would be difficult to obtain this information by any other approach. Estimating subjective probabilities and importance weights are unfamiliar tasks to practicing primary care physicians. The semi-structured interview allowed a flexible situation in which clarification could be provided if the subjects desired. The interview format has the disadvantages of expense, need for interviewer training, and possible interview response effects. However, the trouble and expense was considered justified due to the unfamiliar nature of the questions and the need to ensure that all questions were answered. Training to obtain reliable and objective information from the interview was quite extensive. These training procedures are detailed in Appendix D. The training helped avoid possible interviewer response effects. In addition, interviewers were chosen who were familiar with clinical environments and possessed interviewing skills.

Validity of the Instruments

Ebel (1979) states that validity is not so much a property of a test itself, but of the use made of the results. Valid use of a test is facilitated by stating clearly in the test specifications what it is intended to measure and building it to meet those specifications, checking its reliability, and providing clear instructions on how to use it.

These factors were kept in mind during the construction of the materials used in this study. To address content validity, a panel of experts knowledgeable in the areas of primary care, endocrinology, decision analysis, and measurement research attempted to ensure that the vignettes reflected the management problem of obese female patients. Salient factors from recent literature on utility theory, decision analysis, and obesity management, as well as relevant findings from a recently completed chart audit study addressing the question of obesity referral, were considered in the design of the instruments. Conventionally, validity is demonstrated best by "tryout" after the test has been constructed (Ebel, 1979). This was done in two pilot studies in which five and then three physicians were asked to compare the vignettes to typical experiences with obese female patients in their own practices. On a scale of 1-100, they were asked to evaluate how typical the vignette case information was to the real clinical world. Their average evaluation was 67.

Reliability of the Instruments

Ebel (1979) states that reliability is a necessary condition for instrument quality. Subjects' responses must be consistent to be trustworthy. Test-retest reliability for this study's vignette instrument was examined by replicating six randomly selected cases. The overall test-retest reliability was .86.

Data Collection

Physicians who agreed to participate in the study received a vignette booklet and were asked to complete it prior to the scheduled date for the interview. Subjects who did not complete the booklet prior to the interview completed this task just prior to the interview. The interviewer checked the vignette booklet for completion before commencing the interview. The interview data were collected from subjects in a single 30-60 minute interview conducted by the trained interviewers.

Pilot Studies

"A careful pilot study is the best insurance the research worker has against bias and flaws in design." (Borg and Gall, 1979) Two pilot studies were carried out in the Winter of 1980-81 with these specific objectives in mind:

1. To determine if the procedures would actually produce the data desired and see if quantification and data analysis was possible.

2. To alert the researchers to cues that indicated a rephrasing of questions or revision of recording procedures was required to improve the task.

3. To provide practice and training for study interviewers. Sessions were taped to provide insight into handling questions, to allow sharing of the learning experiences, and to alert the interviewers to potential pitfalls.

A total of eight physicians* participated in the pilot studies: five in the first, three in the second. Physicians gave feedback and suggestions as they progressed through the case vignettes and interview questions. Instruments were refined and ambiguities were clarified based on this feedback. Scaling and editorial changes took place following the first pilot. Two practice cases were added to the case vignette booklet to increase understanding of the task and improve reliability. Contrary to the actual study, interviewers were present as the pilot subjects completed the vignette booklet. Physicians discussed their interpretation of the questions and the reasons behind their responses in a 90-minute session. The taped contents of these sessions were invaluable to the revision of the instruments.

Data Analysis Techniques for the Research Questions

The research questions were as follows:

1. What is the relation between the predictors of physician referral behavior from the model (net SEU's) and number of vignette cases physicians referred? Is it possible to predict the number of vignette cases referred from the scores derived from the model?

---

*Physician subjects were primary care physicians: one specialist, four subspecialists, three residents, three M.D.'s, and two D.O.'s from both academic and community settings.

2. What is the relation between the predictors of physician referral behavior derived from the model and number of vignette cases physicians referred <u>for each specialty</u>? Does the MAUT model fit one specialty group better than another?

3. What is the relation between the predictors of physician referral behavior derived <u>from each attribute</u> and number of vignette cases physicians referred? Is it possible to predict the number of vignette cases referred from the scores derived for each attribute? Of the two attributes, weight reduction and patient satisfaction, is one attribute more highly predictive of number of vignette cases referred than the other?

4. a. Do the predictors of physician referral behavior derived from the model (net SEU's) vary by physician characteristics?

   b. Do the predictors of physician referral behavior derived <u>from each net utility attribute score</u>, net u(WR) and net u(PS), vary by physician groups?

5. What are the properties of the individual attribute parameters (probabilities, importance weights, and values)? Are individual attribute parameters affected by selected physician characteristics (specialty, years in practice) or weight categories (50% and 100% overweight)? What is the relation between individual attribute parameters and number of vignette cases referred?

6. Do the preferred management alternatives, to refer or not to refer, vary by selected categories?

Data analyses were chosen to fit the different purposes of the six research questions.

The purposes of research Questions #1 and #2 were prediction; research Question #3's purpose was largely explanatory. Therefore, the data on these questions was analyzed by linear and multiple regressions. Kerlinger and Pedhazur

(1973) cite the use of regression analysis when the purpose of the research is prediction and explanation.

> Regression analysis can play an important role in predictive and explanatory research framework. In prediction studies the main emphasis is on practical application. On the basis of knowledge of one or more independent variables, the researcher wishes to develop a regression equation to be used for the prediction of a dependent variable, usually some criterion of performance or accomplishment.
>
> In an explanatory framework, on the other hand, the emphasis is on the explanation of the variability of a dependent variable by using information from one or more independent variables.

Research Question #4 inquired about the effects of specialty and number of years of practice on the predictor variables derived from the model. This question was translated into five research hypotheses:

1. Null Hypothesis: With adjustment for years in practice, no difference will be found in the predictors of physician referral behavior derived from the model (net SEU's) based on physician specialty.

2. Null Hypothesis: No difference will be found in the predictors of physician referral behavior derived from the model (net SEU's) based on number of years of practice experience.

These hypotheses were tested for both 50% and 100% overweight patient cohorts by analyzing the data using a one-way Analysis of Covariance (ANCOVA). The analysis of covariance was used in this situation to account for initial differences by specialty in the variable, number of years of practice experience.

3. Null Hypothesis: With adjustment for years in practice, no difference will be found in the predictors of physician referral behavior derived from each attribute, net u(WR) and net u(PS), based on physician specialty.

4. Null Hypothesis: With adjustment for years in practice, no groups by measures interaction in predictors of physician referral behavior derived from each attribute, net u(WR) and net u(PS), will be found.

5. Null Hypothesis: No difference will be found in the predictors of physician referral behavior derived from each attribute, net u(WR) and net u(PS), based on number of years of practice experience.

These hypotheses were tested for both the 50% and 100% overweight patient cohorts by analyzing the data using a one-way Multiple Analysis of Covariance (MANCOVA). This technique was used because of the presence of two dependent variables net u(WR) and net u(PS). As in the ANCOVA with a single dependent variable, the MANCOVA adjusts for initial differences in the variable, number of years of practice experience, among the specialty groups and effectively removes it from consideration.

Research Question #5 examined the descriptive properties of individual parameters (probabilities, importance weights, and values), the effect of selected physician characteristics on these parameters, and the relation between them and the number of vignette cases referred. Here, one-way Analyses of variance, Pearson Product Moment Correlations, and descriptive techniques assessed the data. Research Question #6 examined if the model-derived preferred management alternatives, to refer or not to refer, varied by selected categories. Data relating to this question were analyzed using Chi-squares.

Summary

In this study, 45 voluntary primary care physicians within the specialties of Internal Medicine, Family Practice, and Gynecology responded to a simple linear multi-dimensional utility model on referral management strategies, dealing with the problem of obesity. These responses were compared to referral decisions made to case vignettes describing 100% and 50% overweight female patients. Two data collection formats were used in this study: 1) written structured cases (case vignettes), and 2) a semi-structured interview. Subjects reported the vignettes to be typical of experiences with obese patients ($\bar{X}$ = 67 on a 0-100 scale). Test-retest reliability for the vignette instrument was found to be .86. Extensive interviewer training facilitated the collection of reliable and objective interview data. Data was analyzed in terms of the research questions and hypotheses. Linear regression, one-way ANCOVA and MANCOVA, Chi-Square, and descriptive statistical techniques were used.

CHAPTER IV

PRESENTATION AND ANALYSIS OF THE DATA

Introduction

In this exploratory study, regression, correlational, causal-comparative, and descriptive analytic techniques were used. Data were entered by terminal to the Michigan State University CDC Cyber 750 and were analyzed by SPSS (Statistical Package for the Social Sciences, Nie et al., 1975). Any differences reported to be significant were assessed at the .05 alpha level.

Analyses of Data

Research Question 1

What is the relation between the predictors of physician referral behavior derived from the model (net SEU's) and number of vignette cases physicians referred? Is it possible to predict the number of vignette cases referred from the scores derived from the model?

In order to examine the predictive relation between the primary care physicians' net SEU's and their number of vignette cases referred, a linear regression procedure was used.

For the 50% overweight patients, there was a slight relation between net SEU and number of 50% overweight patients referred ($r = .23$, ns). The mean net SEU50 was entered into the regression equation to examine its contribution to the dependent variable (mean number of 50% overweight cases referred). The mean net SEU50 accounted for only 5% of the variance.

For the 100% overweight patients, the relation between net SEU and number of cases referred was the same ($r = .24$, ns). The mean net SEU100 was entered into the regression equation to examine its contribution to the dependent variable, mean number of 100% overweight cases referred. The mean net SEU100 accounted for only 6% of the variance.

Summaries of the regression analyses for 50% overweight patients and for 100% overweight patients are presented in Table 2.

Table 2

Linear Regression Analyses

a. 50% Overweight Patients

Variable	R Square	Simple r	p
Net SEU50	.05	.23	.13

b. 100% Overweight Patients

Variable	R Square	Simple r	p
Net SEU100	.06	.24	.11

Results suggest that net SEU's were not useful measures for predicting physicians' referral decisions. They account for little of the variance on the dependent variable.

Research Question 2

What is the relation between the predictors of physician referral behavior derived from the model and number of vignette cases physicians referred for each specialty? Does the MAUT model fit one specialty group better than another? Specialty groups were Family Practice, Internal Medicine (General Internists only), and Gynecology.

Three linear regression procedures were used to address this question for both the 50% and 100% overweight patient cohorts. The mean net SEU50 for each specialty was entered into the regression equation to examine its contribution to the dependent variable, mean number of 50% overweight patients referred by that specialty. The procedure was repeated for the mean net SEU100 for each specialty on the dependent variable, mean number of 100% overweight patients referred by that specialty.

Summaries of the regression analyses for 50% overweight patients and 100% overweight patients are presented in Table 3.

Table 3

Linear Regression Analyses for Each Specialty

a. 50% Overweight Patients

Specialty	R Square	Simple r	p
IM (n=15)	.04	+.19	.49
FP (n=15)	.005	-.07	.81
GYN (n=15)	.21	+.46	.09
Overall	.05	+.23	.13

b. 100% Overweight Patients

Specialty	R Square	Simple r	p
IM (n=15)	.002	+.04	.89
FP (n=15)	.14	+.37	.18
GYN (n=15)	.19	+.44	.10
Overall	.06	+.24	.11

Net SEU's were not useful measures for predicting physicians' referral decisions <u>for each specialty</u>. To test whether there were any statistically significant differences in the observed correlations in both sets of results, for the 50% overweight cases, a 95% confidence interval was calculated around the lowest r (Family Practice = -.07). Both r's associated with the Gynecologists and General

Internists fell within this inverval, and as a result are not considered to be from different populations. This could be a function of sampling error due to the small n in each group (15). A larger sample size might show the correlation coefficients to be from different population distributions.

The MAUT model reflects the vignette referral behavior of Family Practitioners and Gynecologists equally for the 100% overweight patients ($r = +.37$ and $+.44$, respectively; $R^2 = .14$ and $.19$, respectively, ns). Again, a 95% confidence interval was calculated around the lowest r (General Internists = +.04). The r's associated with the Gynecologists and Family Practitioners fell within this interval and are not considered different. Again, this finding might be a function of sampling error due to the small n (15), and a larger sample size might show the correlation coefficients to be from different population distributions. Further research is required to substantiate the data trends.

Research Question 3

What is the relation between the predictors of physician referral behavior derived from each attribute and number of vignette cases physicians referred? Is it possible to predict the number of vignette cases referred from the scores derived for each attribute? Of the two attributes, weight reduction (WR) and patient satisfaction (PS), is one attribute more highly predictive of number of vignette cases referred than the other?

In order to examine the predictive relation between the net utility attribute scores, net u(WR) and net u(PS), and the number of vignette cases referred, a stepwise multiple linear regression procedure was used.

For the 50% overweight patients, the two variables analyzed in the step-wise regression were net utility of patient satisfaction, net u(PS50), and net utility of weight reduction, net u(WR50). There was no statistically significant relation between net u(PS50) and number of vignette cases referred ($r = .27$, ns), nor between net u(WR50) and number of vignette cases referred (r = .10, ns). Net u(PS50) entered the equation first and accounted for 7% of the variance on the dependent variable. Adding net u(WR50) to the equation explained no more of the variance.

For the 100% overweight patients, net u(PS100) and net u(WR100) were analyzed by a step-wise regression. There was a low, statistically insignificant relation between net u(WR100) and number of vignette cases referred and net u(PS100) and number of vignette cases referred (r = .20 and .19, respectively, ns). Net u(WR100) entered the equation first and accounted for 4% of the variance on the dependent variable, mean number of 100% overweight vignette patients referred. Adding net u(PS100) to the equation explained an additional 2% of the variance.

Summaries of the multiple linear regression analyses for 50% overweight patients and for 100% overweight patients are presented in Table 4.

Table 4

Multiple Linear Regression Analyses

a. 50% Overweight Patients

Step	Variable	Multiple R	R Square	R Square Change	Simple r	p
1	Net u(PS50)	.27	.07	.07	.27	.08
2	Net u(WR50)	.27	.07	.00	.10	.21

b. 100% Overweight Patients

Step	Variable	Multiple R	R Square	R Square Change	Simple r	p
1	Net u(WR100)	.20	.04	.04	.20	.18
2	Net u(PS100)	.24	.06	.02	.19	.28

In both weight categories, the attributes account for little of the variance on the dependent variable (.07 for the 50% overweight category; .06 for the 100% overweight category). The findings suggest that the net utility attribute scores derived from the model were not useful measures for predicting physicians' patient referral management behavior. However, findings do show that as the degree of overweight increased, the weight reduction attribute gained importance, becoming equal to the patient satisfaction attribute in its relation to the number of vignette cases referred. This is because of the shift in importance weights assigned to the attributes in the 50% overweight and 100% overweight cohorts. This shift is reported later in the results.

Research Question 4a

Do the predictors of physician referral behavior derived from the model (net SEU's) vary by physician characteristics? The characteristics analyzed were specialty and years in practice. The specialties were Family Practice, Internal Medicine (General Internists only), and Gynecology. Years in practice groups were $\leq$ 5 years in practice and > 5 years in practice.

This research question was translated into the following two null hypotheses:

i. With adjustment for years in practice, no difference will be found in the predictors of physician referral behavior derived from the model (net SEU's) based on physician specialty.

ii. No difference will be found in the predictors of physician referral behavior derived from the model (net SEU's) based on number of years of practice experience.

These hypotheses were tested by analyzing the data using two one-way Analyses of Covariance. Covariance was done with net SEU as the dependent variable, years in practice as the covariate, and specialty as the independent variable.

With adjustment for years in practice, the one-way ANCOVAs reflected no statistically significant differences in the physicians' net SEU's based on physician specialty. The number of years in practice as covariate did not vary significantly with net SEU's.

Analysis of Covariance summaries for 50% overweight patients and for 100% overweight patients are presented in Table 5.

Table 5

One-Way Analyses of Covariance for Net SEU's Based on Specialty

a. 50% Overweight Patients

Source of Variation	Sum of Squares	DF	Mean Square	F	p
Covariate (Years in Practice)	.001	1	.001	.115	.74
Main Effects (Specialty)	.021	2	.010	1.038	.36
Explained	.022	3	.007	.731	.54
Residual	.412	41	.010		
Total	.434	44	.010		

b. 100% Overweight Patients

Source of Variation	Sum of Squares	DF	Mean Square	F	p
Covariate (Years in Practice)	.001	1	.001	.102	.75
Main Effects (Specialty)	.001	2	.001	.077	.93
Explained	.002	3	.001	.085	.97
Residual	.279	41	.007		
Total	.281	44	.006		

Research Question 4b

Do the predictors of physician referral behavior derived from each attribute, net u(WR) and net u(PS), vary by physician groups? Physician groups were specialty and years in practice. Specialty groups were Family Practice, Internal Medicine (General Internists only), and Gynecology. Years in practice groups were $\leq$ 5 years in practice and > 5 years in practice.

This research question was translated into the following three null hypotheses:

i. With adjustment for years in practice, no difference will be found in the predictors of physician referral behavior derived from each component attribute, net u(WR) and net u(PS), based on physician specialty.

ii. With adjustment for years in practice, no Groups by Measures interaction in predictors of physician referral behavior derived from each component attribute, net u(WR) and net u(PS), will be found.

iii. No difference will be found in the predictors of physician referral behavior derived from each component attribute, net u(WR) and net u(PS), based on number of years of practice experience.

These hypotheses were answered by analyzing the data using two one-way Multiple Analyses of Covariance (MANCOVA). Covariance was done with net u(WR) and net u(PS) as the dependent variables, years in practice as the covariate, and specialty as the independent variable.

With adjustment for years in practice, the one-way MANCOVAs revealed no statistically significant differences in the physicians' net utility attribute predictor scores (weight reduction and patient satisfaction) based on physician specialty for either the 50% or the 100% overweight patient categories. The WILKS LAMDA Multi-Variate F-test was not significant at the .05 alpha level. There was no Groups by Measures interaction found in net utility attribute scores. The years in practice as a covariate did not vary significantly with net utility attribute scores.

Research Question 5

What are the properties of the individual attribute parameters (probabilities, importance weights, and values)? Are individual attribute parameters affected by selected physician characteristics (specialty, years in practice) or weight categories (50% and 100% overweight)? What is the relation between individual attribute parameters and number of vignette cases referred?

Data for Question 5 were analyzed descriptively and statistically.

1. Probabilities

a. Weight Reduction

Table 6 shows that physicians in general were pessimistic about weight reduction. They indicated that selecting either management alternative, referral or non-referral, made little difference in terms of the probability of weight reduction for either weight category.

Table 6

a. Mean Estimated Probabilities of Weight Reduction in Two Years Without Referral

	50% Overweight	100% Overweight
$P(W_O)$	.63 (.23)	.74 (.21)
$P(W_M)$	.24 (.18)	.22 (.21)
$P(W_I)$	.13 (.12)	.04 (.05)

b. Mean Estimated Probabilities of Weight Reduction in Two Years With Referral

	50% Overweight	100% Overweight
$P(W_O)$	.59 (.25)	.67 (.25)
$P(W_M)$	.26 (.19)	.25 (.23)
$P(W_I)$	.15 (.12)	.07 (.08)

Definitions of Weight Levels:

W_O - Little or no weight reduction
W_M - Weight reduction of approximately one-half the excess weight
W_I - Weight reduction to approximately ideal weight

Note: Standard deviations are in parentheses.

One-way ANOVAs revealed statistically significant differences in the mean estimated probabilities of weight reduction by specialty. These data are reported in Table 11, Appendix E. Using Tukey's method for making post hoc comparisons (Winer, 1971), Family Practitioners were found to be significantly more optimistic than General Internists or Gynecologists, believing that fewer patients would remain at their same weight and expecting more patients to lose about half of the excess weight. T-tests demonstrated no significant differences by years in practice.

Pearson Product Moment Correlations were run between the probabilities of different levels of weight reductions if patient is referred, and the number of vignette cases referred. These correlations were all statistically insignificant for both 50% overweight and 100% overweight patients.

b. Patient Satisfaction

When subjects were asked about the patient's level of satisfaction after two years of management, they indicated approximately 40% of the 50% overweight patients would be satisfied regardless of referral or non-referral. For the 100% overweight patients, referral was expected to increase the probability of patient satisfaction from .29 to .37. A paired t-test found this increase to be significant at the alpha = .05 level (p = .006). These results are shown in detail in Table 7.

Table 7

a. Mean Probabilities of Patient Satisfaction Without Referral

	50% Overweight	100% Overweight
PS	.41 (.25)	.29 (.26)
NS	.59 (.25)	.71 (.26)

b. Mean Probabilities of Patient Satisfaction With Referral

	50% Overweight	100% Overweight
PS	.43 (.27)	.37 (.27)
NS	.57 (.27)	.63 (.27)

Definitions of Patient Satisfaction:

PS - Patient satisfied
NS - Patient not satisfied

Note: Standard deviations are in parentheses.

One-way ANOVAs and t-tests revealed no statistically significant differences in physician estimates of the probability of patient satisfaction by specialty or years in practice.

Pearson Product Moment Correlations were run between the probabilities of patient satisfaction on referral and the number of vignette cases referred. For 50% overweight patients, the correlations was .07, and for 100% overweight patients, it was .20. Both are not statistically significant.

2. Importance Weights

Physicians were asked to assess the relative importance of weight reduction and patient satisfaction as desirable goals of the treatment of obesity.

Physicians on the average gave nearly equal importance to weight reduction and patient satisfaction for 50% overweight patients (.48 to .52). In contrast, for 100% overweight patients, weight reduction was considered markedly higher in importance than patient satisfaction (.68 to .32). Results are shown in Table 8.

Table 8

Mean Normalized Importance Weights

	50% Overweight	100% Overweight
Weight Reduction	.48 (.28)	.68 (.24)
Patient Satisfaction	.52 (.28)	.32 (.24)

Note: Standard deviations are in parentheses.

For 50% overweight patients, 20 (44%) subjects ranked patient satisfaction higher than weight reduction, 17 (38%) ranked weight reduction to be more important, and 8 (18%) stated the attributes were of equal importance.

For the 100% overweight patients, 32 (71%) subjects ranked weight reduction to be more important than patient satisfaction, 6 (13%) felt patient satisfaction was more important, and 7 (16%) ranked the two attributes of equal importance.

Individual physicians were consistent in their ranking order of the attributes, weight reduction and patient satisfaction, across the two weight categories. In particular, 16 of the 17 who ranked weight reduction higher for the 50% overweight patient group also ranked it higher for the 100% overweight patient group. Five of the six who ranked patient satisfaction as most important for the 100% overweight category also ranked this most important in the 50% overweight patient group.

One-way ANOVAs and t-tests revealed no statistically significant differences in the importance weights for either attribute by specialty or years in practice.

Pearson Product Moment Correlations were run between the normalized importance weights of weight reduction and patient satisfaction and the number of vignette cases referred. For the 50% overweight patients, the correlation with the weight reduction importance weight was -.21; the correlation with the patient satisfaction importance weight was +.21. In the 100% overweight cohort, the correlation with the weight reduction importance weight was -.05; the correlation with the patient satisfaction importance weight was +.05. These correlations were all statistically insignificant.

3. Values

Physicians' mean values for losing approximately one-half of the excess weight did not vary from 50% to 100% overweight patients (mean values were .25 (.15) and .27 (.12), respectively).

One-way ANOVAs and t-tests revealed no statistically significant differences in the mean values for losing approximately one-half of the excess weight by specialty or years in practice.

Pearson Product Moment Correlations were run between the values for losing approximately one-half of the excess weight and the number of vignette cases referred. Table 9 shows that the resulting correlations were low.

Table 9

Correlation Between the Values for Losing Approximately One-Half of the Excess Weight and the Number of Vignette Cases Referred

	Number of 50% Overweight Cases Referred	Number of 100% Overweight Cases Referred
Values for Losing Approximately One-Half of the Excess Weight	r = -.31 (p = .02)	r = -.15 (ns)

Research Question 6

Do the preferred management alternatives, to refer or not to refer, vary by selected categories? The categories examined were sex, degree, practice type, practice location, and perception of self as overweight.

According to the model, 24 physicians favored referral, 8 favored non-referral, and 13 indicated indifference to the management choice for the 50% overweight patients. For the group of 100% overweight patients, 30 favored referral, 2 favored non-referral, and 13 indicated indifference. Physicians' responses across the two weight categories again demonstrated consistency. The 13 subjects whose model scores demonstrated indifference in management strategies for 50% overweight patients also demonstrated indifference in the 100% overweight category. The two subjects whose scores favored non-referral for the 100% overweight patient group also favored this strategy for the 50% overweight group.

The Chi-square analyses revealed no statistically significant relation between categorical expected utility (model-derived preferred management alternatives) and selected categories for either the 50% overweight or 100% overweight patients. These data are summarized in Table 10.

Table 10

Physician Characteristics for Favored Management Strategy

a. 50% Overweight Patients

	Sex		Degree		Practice Type		Practice Location		Perception of Self as Overweight
	F	M	DO	MD	Academic/	Community	Lansing	Grand Rapids	
Referral (n = 24)	4	20	6	18	6	18	22	2	11
Indifference (n = 13)	4	9	1	12	5	8	12	1	8
Non-Referral (n = 8)	1	7	3	5	1	7	7	1	6
	9	36	10	35	12	33	41	4	25

Number of Subjects

b. 100% Overweight Patients

	Sex		Degree		Practice Type		Practice Location		Perception of Self as Overweight
	F	M	DO	MD	Academic/	Community	Lansing	Grand Rapids	
Referral (n = 30)	5	25	8	22	7	23	27	3	17
Indifference (n = 13)	4	9	1	12	5	8	12	1	8
Non-Referral (n = 2)	0	2	1	1	0	2	2	0	0
	9	36	10	35	12	33	41	4	25

Number of Subjects

Summary of Findings

Data for the purpose of this study were analyzed by using linear regression, one-way ANCOVA, one-way MANCOVA, one-way ANOVA, Chi-Square, and descriptive techniques.

The relation between the predictors of physician referral behavior derived from the model and number of cases physicians referred in response to a series of case vignettes was explored. There was a slight statistically insignificant relation between the predictors of physician referral behavior derived from the model (net SEU's) and number of cases referred ($r = .23$ and .24 for the 50% and 100% overweight patient cohorts, respectively). Net SEU's were not useful measures for predicting physicians' referral management behavior since they account for little of the variance on the dependent variable, number of vignette cases referred (.05 for the 50% overweight patient cohort; .06 for the 100% overweight patient cohort). Similar findings were found when, for each separate specialty, the relation between the predictors of physician referral behavior derived from the model and number of vignette cases referred was examined.

The relation between the predictors of physician referral behavior derived from each attribute in the two-attribute model and the number of cases physicians referred were explored. Net utility attribute scores derived from the model accounted for little of the variance on the dependent variable, number of vignette cases referred (.07 for the 50% overweight category; .06 for the 100% overweight category). They were not useful measures for predicting physicians' referral management behavior. As the degree of overweight increased, the attribute weight reduction gained importance, becoming equal to the patient satisfaction attribute in its relation to the number of vignette cases referred.

Physician characteristics within the net SEU's; the net utility attribute scores, net u(WR) and net u(PS); and the individual attribute parameters (probabilities, importance weights, and values) were examined. Correlations were run between the individual attribute parameters and number of vignette cases referred. There were no statistically significant differences in the physicians' net SEU's or net utility attribute scores, net u(WR) and net u(PS), based on physician specialty or years in practice. Subjects were more alike than different in their estimates of probabilities, importance weights, and values. Estimates in general varied as much within groups as between groups. Statistically significant differences were not found in the attribute parameters of importance weights or values by specialty or years in practice. No statistically significant differences were found in the patient satisfaction parameter of probability by specialty or years in practice, or in the weight reduction parameter of probability by years in practice. Statistically significant differences were found in the probability of weight reduction parameter by specialty, but not by years in practice. Family Practitioners were found to be significantly more optimistic than General Internists or Gynecologists, believing that fewer patients would remain at the same weight and expecting more patients to lose about half of their excess weight. Pearson Product correlations showed only slight non-significant relations between each of the individual parameters for both attributes and the number of vignette cases referred.

The preferred management alternatives, to refer or not to refer, were examined to determine if they varied by selected categories. The categories examined were sex, degree, practice type, practice location, and perception of self as overweight. Chi-square analyses demonstrated no statistically significant relation between the categorical model-derived preferred management alternatives and any of these selected categories.

CHAPTER V

OVERVIEW; DISCUSSION AND INTERPRETATION OF RESULTS; IMPLICATIONS

Overview

Patient management decisions require clinicians to make complex and critical diagnostic and therapeutic choices based upon what they expect and prefer to happen. Physicians realize that good patient management is based on more than knowledge of scientific facts. They must combine these facts appropriately, balance the dangers and discomforts of a procedure against the value of information to be gained, and recognize the probabilistic nature of outcomes (Donabedian, 1976). In choosing management strategies, physicians commonly agree that there should be a positive outcome for the patient from the selected treatment. The use of outcomes as criteria for assessing quality care is recommended by a number of clinicians (Williamson et al., 1975; Laxdal et al., 1978).

Formal decision analysis involves the application of decision models, such as the multi-attribute expected utility theory model, to problem situations to permit the evaluation of alternative actions. Decision analysis is particularly useful for analyzing complex clinical problems because it is devised to compensate for problem solvers' limitations and biases by assisting in the apprehension, selection, and combination of complex multiple cues. It facilitates the incorporation of relevant variables, values, and uncertainties in a systematic and unbiased manner. A simple multi-attribute expected utility theory model (SMAUT) is considered particularly appropriate for the study of clinical management decisions because its focus is on outcomes. In addition, the explanation of its application and the interpretation of its results can be communicated with relative ease.

The multi-attribute utility theory model has been described as a method for dealing with acute and critical clinical problems such as renal failure, cancer, and burns, but its performance when applied to long-term ambulatory problems has not been investigated. This study examined how a MAUT model performed when applied to the long-term ambulatory problem of obesity. The two attributes (patient goals) chosen to formulate the model were weight reduction and patient satisfaction with management. The model was used to analyze the management decisions of a sample of 45 physicians to refer or not to refer obese female patients to an endocrinologist. The decisions of the subjects generated from the model were compared to intuitive management decisions made by these physicians to a series of case vignettes depicting similar patients. The study subjects were primary care physicians with varied numbers of years of practice experience within the specialties of Internal Medicine, Family Practice, and Gynecology.

Six research questions were asked. The data collected were physician responses to a series of case vignettes and to a semi-structured interview. Linear regression, one-way ANCOVA and MANCOVA, one-way ANOVA, Chi-Square, and descriptive techniques were used to analyze the data. The research questions and results are reported below.

1. What is the relation between the predictors of physician referral behavior from the model (net SEU's) and number of vignette cases referred? Is it possible to predict the number of vignette cases referred from the scores derived from the model?

Results: Net SEU's were not useful measures for predicting physicians' referral management behavior since they accounted for little of the variance on the dependent variable.

2. What is the relation between the predictors of physician referral behavior derived from the model (net SEU's) and number of vignette cases physicians referred for each specialty?

Results: Net SEU's were not useful measures for predicting physicians' referral decisions for each specialty since they accounted for little of the variance on the dependent variable.

3. What is the relation between the predictors of physician referral behavior derived from each attribute and number of vignette cases physicians referred? Is it possible to predict the number of vignette cases referred from the scores derived for each attribute? Of the two attributes, weight reduction (WR) and patient satisfaction (PS), is one attribute more highly predictive of number of vignette cases referred than the other?

Results: Net utility attribute scores derived from the model accounted for little of the variance on the dependent variable. They were not useful measures for predicting referral decisions in the vignettes. As the degree of overweight increased, the weight reduction attribute gained importance, becoming equal to the patient satisfaction attribute in its relation to the number of vignette cases referred.

4. Do the predictors of referral behavior derived from the model, net SEU's, net u(WR), and net u(PS), vary by specialty or years in practice?

Results: There were no statistically significant differences in the physicians' net SEU's or net utility attribute scores, net u(WR) and net u(PS), based on physician specialty or years in practice.

5. What are the properties of the individual attribute parameters (probabilities, importance weights, and values)? Are individual attribute parameters affected by selected physician characteristics (specialty, years in practice) or weight categories (50% and 100% overweight)? What is the relation between individual attribute parameters and number of vignette cases referred?

Results: Subjects were more alike than different in their estimates of probabilities, importance weights, and values. Estimates in general varied as much within groups as between groups. Statistically significant differences were not found in the attribute parameters of importance weights or values by specialty or years in practice. No statistically significant differences were found in the probability parameter for patient satisfaction by specialty or years in practice, or in the probability parameter for weight reduction by years in practice. Statistically significant differences were found in the probability of weight reduction by specialty. Family Practitioners were found to be significantly more optimistic than General Internists or Gynecologists about fewer patients remaining at the same weight and more patients losing about half of their excess weight. Pearson Product Moment Correlations showed only slight non-significant relations between each of the individual parameters for both attributes and the number of vignette cases referred.

6. Do the preferred management alternatives, to refer or not to refer, vary by selected categories? The categories examined were sex, degree, practice type, practice location, and perception of self as overweight.

Results: There were no statistically significant relations between the categorical model-derived preferred management alternatives and any of these selected categories.

Discussion and Interpretation of Study Results

The purpose of MAUT and other decision analytic models is to help decision makers construct explicit models of problem situations which they are accustomed to handling implicitly. Making the decision process explicit can help decision makers and analysts clarify and simplify the management problem and become aware of whether or not the decision is sensitive to factors they thought were important. Shepard (1964) points out that people may use quite different evaluation models than the ones they believe they are using. Johnson and Huber (1977) state that decision makers may not be aware of the attributes that apply to specific situations. Formal decision analysis also allows decision makers to structure the problem situation, i.e., name key decision makers, select decision alternatives, identify attributes that affect the choice of decision alternatives, and specify key uncertainties; and to re-examine this structure following the analysis. This process can lead to the development of decision rules that can be used to organize a large number of similar cases, or it can lead to the conclusion that restructuring the problem situation should be considered. In addition, formalized decision analysis helps decision makers determine if their choices are consistent with their stated preferences and opinions of the uncertainties involved.

A simple MAUT model, a specific type of formal decision analysis, was applied to the problem of obesity. These results were compared to physician responses to case vignettes depicting similar patients. The findings will be discussed in terms of the stated purpose of the MAUT model.

1. Decision makers and analysts can become aware of whether or not the decision is sensitive to factors they thought were important. Management problems can become clarified.

An examination of several of the study results shows that the MAUT model, as formulated, did not describe (predict) what the physicians did in managing the vignette cases. Physician referral behavior in response to the case vignettes was not motivated primarily by the two selected patient goals or attributes, weight reduction and patient satisfaction.

Two explanations can be offered for this result. First, the physicians may have considered one or both of the selected patient goals when making management decisions for the patients depicted in the vignettes, but may have viewed the goal parameters differently. They may not have considered probabilities or importance weights or values, or they may not have combined these elements as the model requires. The literature on judgment suggests human problem solvers have difficulty combining these elements intuitively (Slovic and Lichtenstein, 1971; Weinstein et al., 1980). The second explanation may be that the physicians did not attend to any aspect of the selected outcomes, weight reduction or patient satisfaction with management, when making vignette referral decisions. For the problem of obesity, there may have been other factors more salient than the outcomes selected that influenced the physicians to refer or not to refer vignette patients.

The first explanation will be considered systematically. If physicians intuitively considered only a single attribute or the individual attribute parameters rather than combinations of these elements when responding to the vignettes, the correlation between any of the elements of the model and the vignette referral behavior could have been lost through the arithmetic combination required by the model. That is, although the combined model-derived scores (average net SEU's) accounted for little of the variance on the dependent variable, it is possible that when the relation between either the individual net attribute utilities or individual attribute parameters and the number of vignette cases referred is examined, the correlations would be much higher. However, an examination of the individual net attribute utilities and parameter estimates within each of the attributes and their relation to the number of vignette patients referred ruled out this possibility. For the 50% and 100% overweight cohorts respectively, 93% and 94% of the variance on the dependent variable was still unexplained after the effects of the two net attribute utilities had been accounted for. It appears that neither of the attributes were major factors in the referral behavior. However, one of the individual attribute parameters by itself, i.e., the probabilities or importance weights or values, could have been affecting the number of vignette patients referred. This possibility was also negated when the correlation between each of the parameter estimates and number of vignette cases referred was calculated and found to be low. Therefore, it can be concluded that significant correlations with vignette referral behavior were not lost through combining the parameters or the attributes. It appears that physicians did not attend to any aspect of the selected patient goals, weight reduction or patient satisfaction with management, when making their vignette referral decisions. The second explanation for the results must be accepted. For the problem of obesity, other as yet unidentified factors were more salient than the selected patient goals in determining physician decisions to refer patients in the case vignettes.

When a problem situation is appropriately structured and the model adequately formulated, it is expected that the model will describe behavior. The model's failure to describe referral behavior in this situation suggests the problem formulation should be re-examined.

2. Decision makers and analysts can assess the structure of the decision problem based on the results of the analysis.

Upon reviewing the results and the structure of the decision problem, we may judge that the situation was appropriately formulated, and that the two goals should indeed be the important attributes. Then this prototypical problem situation can be used to develop decision rules that can be applied to a larger number of similar cases. In contrast, re-examination can lead to the conclusion that the decision problem was inappropriately formulated and should be restructured. The discussion addresses each position.

a. The problem situation was appropriately structured and formulated

There is direct and indirect evidence that the decision problem in obesity was appropriately structured in terms of the attributes. In the opinion of a panel of research experts, including two primary care physicians who manage obese female patients, the attributes of weight reduction and patient satisfaction incorporated patient goals that defined successful management of the obese patient. Indirect evidence was provided by the eight pilot study physicians who assessed the effects of the referral alternatives on the probability of achieving weight reduction and patient satisfaction, and assigned importance weights to these attributes. At no time did these physicians suggest that the attributes were inaccurate or insufficient. In addition, the attributes were technically appropriate in that they were limited in number, could be decomposed into levels so that values could be adequately assessed, and were operational, i.e., meaningful, to the physician as well as to the patient. Keeney and Raiffa (1976) report that these are criteria of adequate attributes.

If, upon re-assessing these results, the physicians continue to believe that these attributes should be the factors that determine management choices in obesity, but now know that other factors are directing their referral behavior, they can recognize an inconsistency. This awareness can lead to the development by medical committees of decision rules to apply to similar cases. In this way, the simple MAUT model can serve in a prescriptive fashion, assisting decision makers to overcome the inherent limitations and biases of intuitive decision making so that choices are consistent with beliefs and preferences. Wendt and Vlek (1975) and Marshak (1964) emphasize the prescriptive role for MAUT, noting that intuitive decisions will not highly correlate with model-based decisions in complex environments. The finding that the model-predicted decision did not vary by selected physician characteristics is not unexpected. It is, rather, an example of how decision analysis can aid the decision maker. The model forces the decision maker to focus on particular, explicitly chosen goals and ignore irrelevant variables. Aschenbrenner and Kasubek (1978) and Einhorn (1972) demonstrated that greater inter-decision maker agreement was achieved using the focus provided by the model than when decision making was intuitive.

b. The problem situation is not appropriately structured and formulated

Upon re-examining the problem situation, the clinicians and analysts may conclude that the model failed to describe physician behavior because the problem situation was inadequately formulated. Obesity is a long-term ambulatory health care problem that may have more social sanctions than medical risks. When management involves no real risk or benefit to the patient, a descriptive model may need to incorporate attributes not directly related to the health of the patient. The pessimism about good outcomes for the obesity problem (Penick et al., 1971), the lack of control felt by the physician over patient outcomes, and the minimum benefit or risk associated with most of the treatment alternatives, make the medical benefits of one procedure over another difficult to measure. To formulate a descriptive model for problems like obesity, decision analysts and clinicians may need to consider attributes that reflect patient and physician preferences as well as desired medical outcomes.

The problem situation as structured here did not include direct assessment of patient preferences. Physicians expressed their beliefs about patient satisfaction with management and assigned importance weights to that goal, but patients' opinions were not sought directly. Because the nature of the problem of obesity requires active participation by the patient in its management, patient preferences may greatly influence the physician's choice of strategy. The MAUT model may have been more descriptive if patient preferences had been incorporated into its formulation. McNeil et al. (1978), in a study of patient attitudes, report that patient preferences influence the choice of management strategy.

In situations where there is little medical risk or cost to the patient in a chosen management alternative, such as the referral alternative considered in this study, physicians' professional goals may shift from physiological to psychological considerations. Patients may wish to be referred to a specialist, and physicians, recognizing they have no medical solution, may comply. Rothert et al. (1982), in a study of medical referrals of obese patients to an endocrinologist, found that both patient request for referral and desire of the physician to transfer management were factors in determining referral. Perhaps a shorter time horizon attached to the patient satisfaction attribute might have captured these motives, i.e., the motives might have been captured if physicians had been asked to estimate immediate patient satisfaction with referral, rather than satisfaction with the choice of management after two years.

The proven ability of the MAUT model to describe and predict in other contexts is further supportive evidence that the model may have been inadequately formulated for the obesity problem. Einhorn and McCoach (1977), Gardner and Edwards (1975), Klahr (1969), Stimson (1969), Hoepfl and Huber (1970), Gorry et al. (1973), and Gustafson and Holloway (1975) found simple linear MAUT models to be descriptive in such problem situations as choosing basketball players for the All-Star team, evaluating coastal zone requests for development, deciding on student admissions, accepting public health federal grant applications for funding, assessing the performance of faculty members, evaluating a treatment for renal disease, and developing a burn severity index.

3. Decision makers can determine if their choices are consistent with their stated preferences and opinions of the uncertainties involved.

Data from this study provides some evidence that subjects did not make choices that were consistent with their opinions of the uncertainties and preferences involved. For example, physicians in general were pessimistic about weight reduction for both weight cohorts, regardless of management strategy (referral or non-referral). However, they referred on the average more 100% overweight patients than 50% overweight patients, $\overline{X}$ = 4.5 (2.4) compared with $\overline{X}$ = 3.8 (1.6). This finding suggests that factor(s) other than the probability of weight reduction motivated the choice of referral strategy. Referral strategy also was not significantly associated with the clinicians' stated preferences, i.e., their assigned importance weights or assessed values. The physicians indicated weight reduction was markedly more important for the 100% overweight patient cohort than the 50% overweight patient cohort (mean normalized importance weights being .68 (.24) and .48 (.28), respectively). One might, therefore, expect the physicians to refer more 100% overweight patients. However, the correlation between the importance weight parameter for weight reduction in the 100% overweight group and number of 100% overweight vignette cases referred was -.05. Referring more 100% overweight patients also was not related to the clinicians' mean derived values for losing one-half the excess weight. These values did not vary between the two weight groups. They were .25 (.15) and .27 (.12) for the 50% and 100% overweight cohorts respectively. To summarize, whatever the factors were relating to referral, they were not the importance of weight loss or its probability, or the value of losing one-half the excess weight.

Further, physicians believed only one of the probability parameters would be significantly affected by choosing the referral strategy, i.e., the probability of patient satisfaction for the 100% overweight. They asserted that the choice of referral would increase significantly ($p = .006$) the probability of patient satisfaction; yet the correlation between the probability of patient satisfaction and number of vignette patients who were referred remained low ($r = .20$).

Awareness by an individual physician of inconsistency between preferences and beliefs, on the one hand, and actual behavior on the other, provides an opportunity to reassess the problem and determine if the consistency of individual intuitive decision making can be improved.

In addition to providing insight into the consistencies or inconsistencies of the decision makers' choices, the techniques of decision analysis also help all parties involved in the decision process to recognize variations in their evaluations of the probabilities, importance weights, and values associated with the selected attributes. As stated earlier in this study, the results revealed that the subjects, on the average, were more alike than different in their estimations of probabilities, importance weights, and values. Estimates in general varied as much within groups as between groups. Statistically significant differences were found on only one parameter, the probability of weight reduction by specialty. Family Practitioners were significantly more optimistic than General Internists or Gynecologists, believing that fewer patients would remain at the same weight and expecting more patients to lose about half the excess weight. Across groups, the subjects in general were found to be similar in their estimates of probabilities, importance weights, and values; however, there was individual variation within the parameter estimates. This variation resulted in individual differences in favored model-derived management alternatives. For the 50% overweight cohort, 24 out of 45 physicians favored referral, 8 favored non-referral, and 13 were indifferent to the alternatives. For the 100% overweight cohort, 30 physicians favored referral, 2 favored non-referral, and 13 were indifferent. An awareness of variations in the opinions, preferences, and management choices among practicing and student physicians provides a focus for educational discussions. Feedback from these discussions can make explicit the reasons for these variations. Both effective and inadequate approaches to analyzing complex decisions may become evident through this educational exchange.

Methodological Considerations

Assessing Values

For this study, each subject's value for losing approximately one-half the excess weight, $\bar{v}(W_M)$, was constructed by the researchers. Each physician's estimates of the increase in patient morbidity associated with stated weight levels in the vignettes was used to derive the value. This research-derived method for assessing $\overline{v}(W_M)$ contrasts with the traditional alternative of eliciting the value estimates from each subject. It was an attempt to follow the recommendation of Slovic et al. (1982). "One could hope that further research and analysis would identify better ways to ask questions about values."

The researcher-derived procedure adopted for this study had two major merits. In the opinion of the decision making research team, which included two primary care physicians, the values for $\overline{v}(W_M)$ for each subject could be adequately defined in terms of the physician's estimates of the increase in patient morbidity associated with stated weight levels in the vignettes. Estimating the degree of patient morbidity is a familiar task for a physician; obtaining an accurate value response for each subject was thereby facilitated. In addition, the subjects were not required to take additional time to learn a value assessment technique, i.e., category scaling, standard gamble, or time trade-off, because they had already provided morbidity estimates in completing the vignette booklets. Thus, the procedure adopted avoided an overly long data collection procedure which might have decreased the quality of the data.

Utilizing a research-derived approach receives support from the literature. The traditional methods of preference assessment (category scaling, standard gamble, and time trade-off) are being questioned on their capacity to assess the decision maker's avowed preferences since responses appear to vary according to the method selected (Slovic et al., 1982). The use of more than one traditional preference elicitation method within a particular study is recommended to provide evidence of the validity of the value responses (Hershey et al., 1981). In addition, each of the traditional preference elicitation methods presents difficulties when the problem situation involves long-term health states. The standard category scaling technique is considered inappropriate because physicians are not accustomed to assessing values of alternate health states on a utility scale using arbitrary units. When decision makers are presented with unique and unfamiliar tasks, it may be difficult for them to express their avowed values (Slovic et al., 1982). Considerable time is required to train interviewers and subjects in the use of the standard gamble method. As well, the data collection procedure to obtain the value estimations is relatively lengthy. The gamble method is also based upon a set of fundamental axioms of which the construct validity is now being challenged (Eraker and Politser, 1981; Kahneman and Tversky, 1979; Hershey et al., 1981). A time trade-off technique is not appropriate for conditions such as obesity because the attribute levels and resultant outcomes cannot be defined in terms of trading off a shorter healthy life for a longer life with impaired health. Data collection and training periods also are rather lengthy.

It is of interest to note the values obtained by the research-derived method. The mean value for losing one-half of the excess weight for the 50% and 100% overweight groups were .25 (.15) and .27 (.12), respectively. It was expected that the values for losing one-half the excess weight would differ more for the two weight cohorts since in one case, the patient is down to 25% above ideal and in the other, the patient is still 50% overweight. Further investigation is required to determine in this situation if the physicians were accurately describing their indifference to a 25% weight differential, or if the scale was insensitive to differences which exist.

Measurement of Clinical Behavior

Case vignettes were used to depict clinical referral behavior as standardized cases control for the confounding variable of patient differences. Further, a large number of clinical cases can be sampled in a limited time. When asked to evaluate how typical the vignette cases were to the real clinical world, the physicians' average evaluation was 67 on a 0-100 scale, indicating the face validity of the cases was satisfactory. Retrospectively, however, the external validity of the vignettes may be questioned. Patient request for referral was systematically incorporated into eight cases. Physicians reported that this factor occurred more frequently in the vignettes than in their practice. Consideration must be given to the possibility that this non-representative factor may have biased physician referral behavior in response to the vignettes away from practice behavior. It could be that model-derived behavior may be more highly related to actual practice behavior than to vignette behavior. Although further investigations are required to test this hypothesis, a chart audit study by Ravitch et al. (1982) does suggest referral to an endocrinologist occurs less frequently in reality than occurred in the vignettes. Their study estimated the referral rate of clinic-seen obese patients to endocrinologists to be 1%. In this study, across all subjects the number of vignette cases referred ranged from 0-19, with a mean of 8. The cases in which the patient requested referral to an endocrinologist were the most frequently referred.

Implications

Research

To gain a further understanding of the management of obesity, research is recommended replicating this study's methods, but restructuring the problem situation by incorporating non-medical attributes, patient preferences, and physician goals, as well as considering different management alternatives. With these restructured problem situations, MAUT models could be applied to the obesity management problem to see if the model more accurately describes behavior.

It would be useful to know if the nature of long-term clinical management problems per se affects the relation between clinical behavior derived from a MAUT decision aid and actual clinical behavior, or if obesity is a special case.

Would MAUT fail to describe management behavior if the model were applied to such long-term management problems as hypertension and diabetes, where management choices do have impact on patient outcome and present risks and benefits to the patient?

Techniques used to derive the probabilities, importance weights, and values appeared effective. The subjects did not express difficulty in providing subjective probabilities or estimating importance weights. The estimates were relatively consistent with what one would expect across subjects and weight categories. Further research is recommended, however, to determine whether researcher-derived methods for assessing values are consistent with the declared values of decision makers.

Dawes and Corrigan (1974) and Edwards (1977) attest to the robustness of the simple multi-attribute utility model. The model does, however, contain a number of assumptions (Einhorn and McCoach, 1977), i.e., utility functions for the attributes are linear, the total utility of an outcome is an additive function of the utilities of the attribute making up the outcome, levels of attributes are monotonic (more is always better than less or less is always better than more), and attributes are valued independently of one another. This study assumed that the assumptions of the model had been met. Edwards (1977) reports that "quite substantial amounts of deviation from value independence will make little difference to the ultimate number U_i." However, further studies could be carried out to test the linearity and independence assumptions.

The problem situation of obesity with the chosen management alternatives could be re-examined by the application of process-oriented models of human choice behavior, such as the Lens Model or the Information Processing Model. The findings from those studies could be compared with these results, which are based on the application of the goal-oriented MAUT model.

In future, the clinical choices based on MAUT models need to be compared to choices in actual clinical settings, i.e., hospitals, clinics, or physicians' offices. What proportion of a group of 24 obese patients seen by a primary care physician are referred to an endocrinologist; what proportion are managed by the practitioner? How does actual practice behavior relate to the behavior one would expect from using the model?

Principles of decision analysis could be incorporated into the assessment and teaching of medical care. Decision analysts and clinicians could collaborate in the design of decision analytic teaching and evaluation tools (e.g., health records) which could capture information such as probabilities, preferences, and benefits and risks. Such tools do not exist at present. Research could then be carried out comparing present methods of assessing and teaching medical care with those using principles of decision analysis. That is, in comparable settings, non-randomized control trials could be carried out to compare decision-analytically designed tools and present tools in their capacity to facilitate quality patient care, medical education, and hospital policy assessment. For similar cases, objective information such as number and costs of diagnostic tests, days stay, patient complications, condition on discharge, and re-admission required could be collected and compared from new and old tools. Subjective opinions from experts, teachers, and students could be sought to compare the educational effectiveness of revised tools with the current methods.

Further research is advisable to determine when the limitations of human processing abilities interfere with quality care so that refined decision aids are appropriate and practical, and also to define the degree of sophistication of a decision aid that is appropriate and practical (Fischhoff, 1977). The formal properties of the decision situation may dictate when it is advisable to use specific decision analysis formats. For example, for some decision problems, it may be more appropriate to follow simple decision analysis guidelines rather than to carry out all the computations required by a model's full application. The guidelines could be formulated as heuristic principles, for example, "when exercising clinical judgment, consider the probabilities of the relevant important outcomes and outcome preferences (importance weights and values) of the key decision makers." Such heuristics may be useful aids to efficient clinical choices.

Quality Care, Medical Education, and Health Policy

The principles of decision analysis bring a new perspective to planning, conducting, and assessing quality care, medical education, and health policy programs. Health care delivery can be viewed in terms of its consistency with the stated preferences and opinions of the particular key decision makers, i.e., physicians, patients, or society, rather than in terms of consistency with standardized protocols. One can study and discuss the rationale behind decisions, not just the behavior itself. Decision analysis provides the means to assess and teach medical care in terms of patient goals or outcomes, and helps to explicate the benefits and risks to the patient associated with the management alternatives under considerations.

The principles of decision analysis have practical implications for several areas within quality care, medical education, and health policy.

1. Quality Care Committees

A new approach to quality care assessment could be offered to Quality Care Committees. Clinicians and decision analysts could form quality care and medical education committees to structure clinical problems. Physicians would identify appropriate clinical problems for study, identify the attributes or goals for the specified problem, and specify alternative choices of action. Decision analysts working with clinicians would take the responsibility for determining which decision analytic procedures are best suited to operationalize the application.

They would also take the primary responsibility for the decision analytic computations. The entire committee would determine which measures of clinical behavior should be used to compare with the results from decision analysis; for example, actual clinical behavior, simulated patients, hypothetical cases, medical records, computerized simulations, etc. They would determine when the formal decision model should be applied and when it is practical to use some of the component ideas.

2. Quality Care Educational Programs

Quality care educational programs could be designed to address a new type of identified educational need, i.e., reasoning inconsistencies or deficiencies. Quality care and educational committees could organize and implement educational and quality care programs based on discussions explicating the reasons for variations within the clinical decision makers' beliefs and preferences. When deficiencies in reasoning are identified and attributed to knowledge or performance gaps, appropriate educational expertise and interventions could be sought.

3. Clinical Protocols Based on Patient Outcome

Discussing the adequacy of a problem structure can lead to efficient clinical protocols based on patient outcome. If problem situations are appropriately structured, satisfactory decision rules to guide physicians and student physicians can be developed. Edwards (1977) states, "decision rules can provide an ordered way to proceed from information and data to values and decisions." They allow the decision makers to focus upon relevant issues, rather than on aspects of the situation that most strongly engage their biases. Problems arising in similar situations can be handled more simply, efficiently, and cost-effectively by following such decision rules. Williamson et al. (1975) state that "systematic consideration of relationships between medical care and outcomes can provide a crucial means of cutting through the enormous number of irrelevant variables so often included in the assessment of quality care and continuing medical education."

4. Validation for Clinical Decisions and Health Policy

Validation could be attained for clinical decisions and health policy based on patient benefit and risk. By explicating, through formal decision analysis, the benefits and risks to the patients associated with choosing diagnostic or therapeutic strategies, clinicians and health policy analysts can evaluate the contributions of particular clinical choices, such as referral or laboratory or radiological examinations, to a defined clinical problem or patient, and clarify the value of individual investigations or alternatives. Because health resources are limited, attempts can be made to gain the greatest health benefits for patient and society at the least cost. Decision analysis can provide a rationale for choosing a particular diagnostic or management action for a particular clinical problem or individual patient's case. As treatments become increasingly expensive and potentially invasive, clinical decisions require additional analytic validation (Knowles, 1977). Applications of decision analysis provide a mechanism for this validation.

5. Curriculum

Medical school curriculum committees could investigate the incorporation of decision analysis principles into the curriculum. These ideas, and materials based on them, can be extended to teaching hospitals and clinics to use with undergraduates, graduates, and practicing clinicians as they solve problems and make critical and costly decisions. Taylor (1976) states that formal decision analysis can provide a "different, but complementary approach to teaching in wards and clinics where the familiar approach based on pathology or direct diagnosis is complemented by an operational one which looks at critical decisions made about the patient's management."

6. Individual Assessment

Decision analysis provides an opportunity for individual clinicians to participate in self-education and self-evaluation by examining their own clinical decision making processes and clinical choices. By reflecting upon their decisions in terms of decision analysis principles, individuals could determine if their choices are consistent with the stated preferences and the uncertainties perceived in the situation, and thereby could improve the consistency of their own intuitive decisions.

Summary

This study examined how a simple MAUT model performed when applied to the long-term ambulatory problem of obesity. The model was used to analyze the decision to refer or not to refer obese female patients to an endocrinologist in a sample of 45 primary care physicians. The decisions of the subjects derived from their responses to the model were compared to intuitive management decisions made on a series of case vignettes depicting similar patients.

Six research questions were asked in the study. The data obtained were physician responses to a series of case vignettes and to a semi-structured interview. Linear regression, one-way ANCOVA and MANCOVA, one-way ANOVA, Chi-Square, and descriptive techniques were used to analyze the data. Two major findings are reported:

1. The predictors of physician referral behavior derived from the model did not predict referral behavior in the series of case vignettes.

2. The predictors of physician referral behavior derived from the model did not vary significantly by selected physician characteristics.

The interpretation of the findings have important implications for research, decision analysis, quality patient care, medical education, and health policy.

APPENDICES

APPENDIX A

Preference Assessment Techniques

APPENDIX A

Preference Assessment Techniques

Standard Gamble Technique

The standard gamble technique (Weinstein et al., 1980) requires the decision maker to rank all aspects of the attributes on a utility scale, in this case arbitrarily anchoring the least preferred aspect at 0 and the most preferred aspect at 1. The intermediate attribute levels, however, are not assigned directly as in category scaling, but rather are derived by a gamble method. For example, using the attribute "weight reduction," the evaluator is asked to choose between two alternatives--a gamble between the chance of reaching ideal weight versus a chance of losing no weight, or a guaranteed loss of some weight. Probabilities assigned to the best and worse outcomes of the gamble are changed until, at some probability, the decision maker cannot decide between the gamble and the guaranteed alternative. The probability of the best outcome at this "indifference" point becomes the value of the guaranteed alternative (or intermediate attribute level). In other words, if the decision maker is indifferent between the gamble of 20% chance of reaching ideal weight versus 80% chance of losing no weight, and the guaranteed alternative of losing half the excess weight, the worth of the alternative would be 20%. The decision maker assigns importance weights to each attribute in a ratio to the least important attribute so that the value of levels within attributes are comparable across attributes. Again, as in category scaling, importance weights for each attribute are normalized by summing them and dividing each weight by the total sum. Preference (utility) for a particular attribute level is calculated by multiplying the values elicited for the attribute level by the importance weight placed on each attribute.

The standard gamble or lottery method seems inappropriate to some decision makers. Clinicians are not accustomed to thinking of the value of alternate states of health in terms of gambles, even though clinical problem solving almost invariably requires actions based on an assessment of probabilities. Another criticism of the method concerns the set of axioms upon which it is based (Schoemaker, 1980). The construct validity of these fundamental axioms are now being challenged because of the context boundedness of an individual's value estimations in any particular situation (Eraker and Politser, 1981; Kahneman and Tversky, 1979; Hershey et al., 1981). An advantage of this method is that the values obtained are in terms of probability and therefore do not require re-scaling to fit the decision analysis framework.

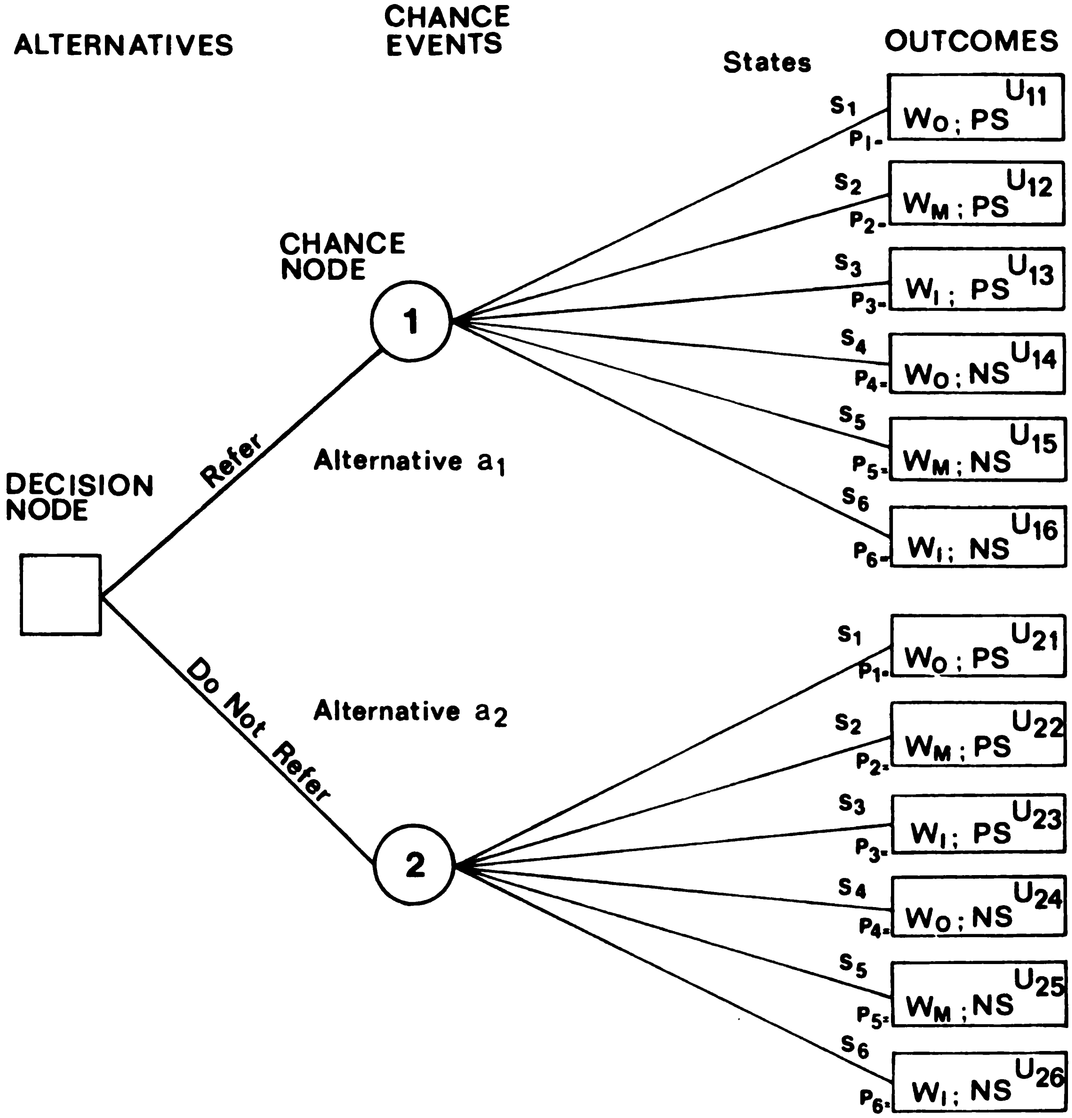
Time Trade-off Technique

In the time trade-off method (Torrance, 1972), decision makers are asked to choose between a long life with impaired health and a shorter life with perfect health. A series of progressively shorter life spans with optimal health are offered as an alternative to a life span of fixed duration with impaired health. The point at which the decision maker cannot choose between the shorter healthy life and the longer life with impaired health is called the "indifference" point, and is expressed as a ratio of years of perfect health to years of impaired health. The indifference point or ratio is considered to be the value of years of impaired health. Since all outcome attributes are expressed in units of time, attribute categories can be directly compared. This method resembles the standard gamble in that the subject is forced to choose between alternatives. The time trade-off method can also be viewed as a special case of category scaling in which the scale values are units of time. People have a greater appreciation of time than they do of probability in general, so the units of time may have greater reality for decision makers than the arbitrary units used in category scaling or the probabilities on bets employed in standard gamble.

APPENDIX B

Decision Tree Representation
of the Structured Problem

APPENDIX B



Chance Node 1 = $P_1 \cdot U_{11} + P_2 \cdot U_{12} + P_3 \cdot U_{13} + P_4 \cdot U_{14} + P_5 \cdot U_{15} + P_6 \cdot U_{16}$

Chance Node 2 = $P_1 \cdot U_{21} + P_2 \cdot U_{22} + P_3 \cdot U_{23} + P_4 \cdot U_{24} + P_5 \cdot U_{25} + P_6 \cdot U_{26}$

KEY

S - State
P - Probability
U - Utility (Preference) = Importance Weight × Value
W_O - Little weight reduction; if any
W_M = Weight reduction down to approximately one-half of excess weight
W_I = Weight reduction down to approximately 10-15% above ideal weight
PS - Patient Satisfied
NS - Patient Not Satisfied

Figure 1.
THE DECISION TREE REPRESENTATION OF THE STRUCTURED PROBLEM

APPENDIX C

Case Vignette Task

Explanation, Directions, Examples of Vignettes

APPENDIX C

Case Vignette Task

Explanation

Physicians see many obese patients, and selecting a management program is sometimes difficult. A group of faculty from the College of Human Medicine at Michigan State University is studying the decisions physicians make related to choice of management for obese patients. This is a study of the basis on which decisions are made, not an attempt to evaluate the appropriateness or correctness of clinical judgment.

We are interested in the management choices you would make in the following hypothetical cases. We particularly wish to have your response because your experience as a practicing physician will contribute significantly toward understanding the basis on which practicing physicians make decisions. Please respond to the cases by thinking about what you would actually do in each situation.

The 30 hypothetical cases presented are very similar in construction. Four identical questions follow each vignette. We recognize that the task may seem somewhat repetitive, but we urge you to complete all of the cases as complete data is necessary for analysis. A break following the completion of 12-15 vignettes may make the task easier.

Your time and cooperation are greatly appreciated.

Directions

PLEASE REMOVE THIS PAGE FOR EASY REFERENCE

There are 30 hypothetical case vignettes on the following pages. You will be asked to respond to questions 1, 2, and 3 by using a 7-point scale. The scales for each of these questions are defined as follows:

Question 1. Given the facts of this case, how likely are you to refer this patient to an endocrinologist?

1. Virtually certain not to refer (0-5%)
2. Fairly certain not to refer (6-20%)
3. Probably would not refer (21-40%)
4. May or may not refer (41-60%)
5. Probably would refer (61-80%)
6. Fairly certain to refer (81-95%)
7. Virtually certain to refer (96-100%)

Question 2. How does this patient's obesity increase her chances of morbidity?

1. Virtually certain not to increase (0-5%)
2. Fairly certain not to increase (6-20%)
3. Probably would not increase (21-40%)
4. May or may not increase (41-60%)
5. Probably would increase (61-80%)
6. Fairly certain to increase (81-95%)
7. Virtually certain to increase (96-100%)

Question 3. Given what you know about this patient, what are the chances that an endocrine disorder is involved in her weight problem?

1. Virtually certain it is not involved (0-5%)
2. Fairly certain it is not involved (6-20%)
3. Probably is not involved (21-40%)
4. May or may not be involved (41-60%)
5. Probably is involved (61-80%)
6. Fairly certain it is involved (81-95%)
7. Virtually certain it is involved (96-100%)

Background Information

The following information applies to each case. Please read it carefully and refer back to it as needed when responding to the cases.

The hypothetical case studies in the vignette booklet concern obese* female patients who come to you as new patients and request a Pap test.

Please consider that each women in these cases is a new patient who comess for her appointment in the middle of the afternoon. After doing your routine history and physical, including the Pap test, you have some time left. You may elect not to use it.

In the process of the history and physical, you and the patient have discussed her weight. The patient's opinions expressed have been consistent throughout the interaction and are those held at the conclusion of her office visit.

An endocrinologist who routinely manages obese patients and a clinical laboratory able to perform all specialized procedures are available.

You may assume that each patient:

1. Comes to your office in no apparent distress.
2. Is able to function and carry on normal daily activities.
3. Engages in no exercise program other than the motor activity inherent in normal daily activities.
4. Has no major clinical problems other than what is implied by the case data.
5. Is covered by health insurance.
6. Has been overweight since her late 20's.

The fourth question following each vignette will ask you to indicate your level of difficulty in responding to the case. We will discuss some of these cases with you later.

---

*The terms "obese" and "overweight" are synonymous in this study. Percent overweight is defined using the Metropolitan Life Insurance tables for ideal weight by height.

Vignette #1

Tina U. is a 45-year-old female who weighs 181 pounds and is 5'6" tall (33% overweight). When you inquire about her eating habits and weight, she states she has been calorie counting for years and knows her intake is between 1000-1200 calories per day. She asks you what you can do about her obesity. Physical examination reveals a patient who is minimally overweight and whose abdomen is free of striae or scars. The remainder of your physical examination is not remarkable.

1. Given the facts of this case, how likely are you to refer this patient to an endocrinologist? (Circle the appropriate number)

   1 2 3 4 5 6 7

   If you circled 5, 6, or 7, what is your major reason for referral?

   _______ Diagnosis
   _______ Take over management of the patient
   _______ Support and reassurance
   _______ Other ________________________

2. How does this patient's obesity increase her chances of morbidity?

   1 2 3 4 5 6 7

3. Given what you know about this patient, what are the chances that an endocrine disorder is involved in her weight problem?

   1 2 3 4 5 6 7

4. Please check the appropriate category indicating the level of difficulty responding to this case:

   _______ E (particularly easy)
   _______ A (average)
   _______ B (particularly difficult)

Vignette #2

Elizabeth Y. is a 42-year-old female who weighs 208 pounds and is 5'4" tall (60% overweight). She states that she does not overeat and believes she has a hormone problem. She asks to be referred to a "hormone problem" specialist. Upon continued questioning, she insists her diet does not exceed 1200 calories per day and strongly repeats her request to see a hormone doctor. On physical examination, you notice that the patient is moderately overweight and that she has a few red striae over her abdomen. The remainder of your physical examination is normal.

1. Given the facts of this case, how likely are you to refer this patient to an endocrinologist? (Circle the appropriate number)

   1 2 3 4 5 6 7

   If you circled 5, 6, or 7, what is your major reason for referral?

   _______ Diagnosis
   _______ Take over management of the patient
   _______ Support and reassurance
   _______ Other ________________________

2. How does this patient's obesity increase her chances of morbidity?

   1 2 3 4 5 6 7

3. Given what you know about this patient, what are the chances that an endocrine disorder is involved in her weight problem?

   1 2 3 4 5 6 7

4. Please check the appropriate category indicating the level of difficulty responding to this case:

   _______ E (particularly easy)
   _______ A (average)
   _______ B (particularly difficult)

APPENDIX D

Interviewer Training

APPENDIX D

Interviewer Training

Prior to Pilot Study I

Interviewer:

The researcher: Ph.D. candidate, Health Record Administrator

Training prior to Pilot Study I consisted of:

a. involvement in the development and revision of instruments,

b. involvement in NLM research meetings for one-year prior to initiation of instruments--discussions involved variables being studied, logistics, and safeguards of interviewer conditions, etc.,

c. study of instruments, and

d. ten years health science background involving working with and interviewing physicians in clinical and research environments.

The Program Associate also participated as an interviewer in Pilot Study I.

Prior to Pilot Study II

Three interviewers: the researcher and two additional Ph.D. candidates:

1. Nurse - with extensive experience working and dialoguing with physicians and previous exposure to and study of Cognitive Processes, including Decision Analysis.

2. Sociology major - with specific experience and background in interviewing skills with previous study in Educational Psychology.

Training prior to Pilot Study II for three interviewers consisted of:

a. study of the Interview Guide,

b. involvement in NLM research meetings discussing the study of Pilot Study I results,

c. analysis and discussion of tapes from Pilot Study I, and

d. conduction of practice interviews, as interviewers and interviewees.

APPENDIX E

Table 11

Mean Probabilities of Weight Reduction by Specialty For Both 50% and 100% Overweight Patients

APPENDIX E

Table 11

Mean Probabilities of Weight Reduction by Specialty For Both 50% and 100% Overweight Patients

		Without Referral				With Referral			
			50%		100%		50%		100%
W_O	IM	$p = .03$	.75 (.19)	$p = .0006$	.83 (.15)	NS	.68 (.28)	$p = .01$	.77 (.22)
	FP		.53 (.23)		.57 (.26)		.49 (.22)		.52 (.28)
	GYN		.61 (.23)		.81 (.08)		.61 (.24)		.73 (.17)
W_M	IM	$p = .02$	.18 (.14)	$p = .003$	.13 (.12)	$p = .01$	.21 (.18)	$p = .004$	.16 (.15)
	FP		.35 (.23)		.38 (.26)		.37 (.22)		.40 (.28)
	GYN		.20 (.10)		.15 (.09)		.19 (.11)		.20 (.15)
W_I	IM	$p = .03$	.08 (.09)	NS	.04 (.04)	NS	.11 (.11)	NS	.07 (.08)
	FP		.12 (.07)		.05 (.05)		.14 (.06)		.08 (.07)
	GYN		.19 (.15)		.05 (.07)		.20 (.16)		.08 (.09)

Standard deviations are in parentheses.

LIST OF REFERENCES

LIST OF REFERENCES

Abraham, S., and Johnson, C. L. Advancedata. *Vital & Health Statistics*, National Center for Health Statistics, U.S. Department of Health, Education and Welfare. Public Health Service Office of Health Research, Statistics, and Technology, No. 51, August 30, 1979.

Abramson, E. E. Behavioral approaches to weight control: An updated review. *Behavioral Research and Therapy*, 1977, *15*, 355-363.

Aschenbrenner, K. M., and Kasubek, W. Challenging the Cushing Syndrome: Multi-attribute evaluation of cortison drugs. *Organizational Behavior and Human Performance*, 1978, *22*, 216-234.

Beach, B. H. Expert judgment about uncertainty: Bayesian decision making in realistic settings. *Organizational Behavior and Human Performance*, 1975, *14*, 10-59.

Bernoulli, D. /Specimen theoriac novae de menoura sortis./ *Commentarie Academic Scientiarum Imperialis Petropolitanae*, 1738, *5*. (L. Sommers in *Econometrics*, 1954, *22*, 22-36.)

Bierman, H., Bonini, C. P., Fouraker, L. E., and Jacdicke, R. K. *Quantitative analysis for business decisions* (Rev. ed.). Homewood, Illinois: Richard D. Irwin, Inc., 1965.

Borg, W. R., and Gall, M. D. *Educational research: An introduction* (3rd ed.). New York: Longman, Inc., 1979.

Bray, G. A. Clinical management of the obese adult. *Postgraduate Medicine*, 1972, *51*, 125-130.

Bray, G. A. The overweight patient. In G. H. Stallerman (Ed.), *Advances in internal medicine*. Chicago: Yearbook Medical Publishers, 1976.

Bunch, W. H., and Andrew, G. M. Use of decision theory in treatment selection. *Clinical Orthopedics and Related Research*, 1971, *80*, 39-52.

Childs, A. W., and Hunter, E. D. Non-medical factors influencing use of diagnostic x-ray by physicians. *Medical Care*, 1972, *10*(4), 323-335.

Collingwood, T. R., and Willet, L. The effects of physical training upon self-concept and body attitude. *Journal of Clinical Psychology*, 1971, *27*, 411-412.

Dawes, R. N., and Corrigan, B. Linear models in decision making. *Psychological Bulletin*, 1974, *81*, 95-106.

Detmer, D. E., Fryback, D. G., and Gassner, K. Heuristics and biases in medical decision making. Unpublished paper, University of Wisconsin, 1977.

Donabedian, A. Measuring and evaluating hospital and medical care. Bulletin New York Academy of Medicine, 1976, 52(1), 51-59.

Drenick, E. J., Bale, G. S., Seltzer, F., and Johnson, D. G. Excessive mortality and causes of death in morbidly obese men. Journal of American Medical Association, 1980.

Ebel, R. L. Essentials of Educational measurement (3rd ed.). Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1979.

Edwards, W. How to use multi-attribute utility measurement for social decisions. IEEE Transactions on Systems, Man and Cybernetics, 1977, SMC-7(5).

Einhorn, H. J. Expert measurement and mechanical combination. Organizational Behavior and Human Performance, 1972, 7, 86-106.

Einhorn, H. J., and Hogarth, R. M. Unit weighting schemes for decision making. Organizational Behavior and Human Performance, 1975, 13, 171-192.

Einhorn, H. J., and McCoach, W. A simple multi-attribute procedure for evaluation. Behavioral Science, 1977, 22, 270-282.

Eisenberg, J. M. Sociological influences on decision-making by clinicians. Annals of Internal Medicine, 1979, 90(6), 957-964.

Eisenberg, J. M., and Nicklin, D. Use of diagnostic services by physicians in community practice. Medical Care, 1981, 19,(3), 297-309.

Eraker, S. A., and Polister, P. Patient participation in clinical decision making. Unpublished paper, 1981.

Fischhoff, B. Decision analysis: Clinical art or clinical science? Paper prepared for the Sixth Research Conference on Subjective Probability, Utility, and Decision Making, Warsaw, September 6-9, 1977.

Fryback, D. G. Use of radiologists' subjective probability estimates in a medical decision making problem. Unpublished doctoral dissertation, University of Michigan, 1974.

Gardner, P. C., and Edwards, W. Public values, multi-attribute utility measurement in social decision making. In M. F. Kaplan and S. Schwartz (Eds.), Human judgment and decision processes. New York: Academic Press, 1975.

Giaugue, W. C. Prevention and treatment of streptococcal sore throat rheumatic fever - a decision theoretic approach. Unpublished doctoral dissertation, Harvard University, 1972.

Ginsberg, A. H., and Offensend, F. L. An application of decision theory to a medical diagnosis-treatment problem. IEEE Transactions on Systems, Science and Cybernetics, 1968, SSC-4,(3), 355-362.

Goldberg, L. R. Simple models or simple processes? Some research on clinical judgments. *American Psychologist*, 1968, *23*, 483-496.

Gorry, G. A., Kassirer, J. P., Essig, A., and Schwartz, W. B. Decision analysis as the basis for computer-aided management of acute renal failure. *The American Journal of Medicine*, 1973, *55*, 473-484.

Grayson, C. J., Jr. *Decisions under uncertainty - drilling decisions by oil and gas operations*. Division of Research, Graduate School of Business, Harvard University, Boston, 1960.

Gustafson, D. H. *Comparison of methodologies for predicting and explaining hospital length of stay*. Unpublished doctoral dissertation, University of Michigan, 1966.

Gustafson, D. H., and Holloway, D. C. A decision theory approach to measuring severity in illness. *Health Services Research*, 1975, 97-106.

Hammond, K. R., and Joyce, C. B. Psychological influences on human judgment, especially of adverse reactions. In F. H. Gross and W. W. Inman (Eds.), *Drug monitoring*. London Academic Press, 1977.

Henschke, U. K., and Flehinger, B. J. Decision theory in cancer therapy. *Cancer*, 1967, *20*, 1819-1826.

Hershey, J. C., Kunreuther, H. C., and Schoemaker, P. H. *Sources of bias in assessment procedures for utility functions*. (In press)

Hoepfl, R. T., and Huber, G. P. A study of self-explicated utility models. *Behavioral Science*, 1970, *15*,(5), 408-414.

Hogarth, R. M. *Judgment and choice*. New York: John Wiley & Sons, 1980.

Howard, R. The foundations of decision analysis. *IEEE Transactions on Systems, Science and Cybernetics*, 1968, *SSC-4*(3), 211-219.

Howard, R. A. An assessment of decision analysis. *Operations Research*, 1980, *28*(1), 4-27.

Huber, G. P. Multi-attribute models: A review of the field and field-like studies. *Management Science*, 1974, *20*, 1393-1402.

Huber, G. P., Daneshgar, R., and Ford, D. L. An empirical comparison of five utility models for predicting job preferences. *Organizational and Human Performance*, 1971, *6*, 267-282.

Johnson, E. M., and Huber, G. P. The technology of utility assessment. *IEEE Transactions on Systems, Man and Cybernetics*, 1977, *SMC-7*, 311-325.

Kahneman, D., and Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrics*, 1979, *47*, 263-291.

Kassirer, J. P. The principles of clinical decision making: An introduction to decision analysis. *The Yale Journal of Biology and Medicine*, 1976, *49*, 149-164.

Keeney, R. L., and Raiffa, H. A critique of formal analysis in public decision making. In A. W. Drake, R. L. Keeney, and P. M. Morse (Eds.), *Analysis of public systems*. Cambridge: MIT Press, 1972.

Keeney, R. L., and Raiffa, H. *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: John Wiley & Sons, 1976.

Kerlinger, F. N., and Pedhazur, E. J. *Multiple regression in behavioral research*. New York: Holt, Rinehart, & Winston, 1973.

Keys, A. Obesity and heart disease (editorial). *Journal of Chronic Diseases*, 1955, *1*, 456-461.

Klahr, D. Decision making in a complex environment. The use of similarity judgments to predict preferences. *Management Science*, 1969, *15*(11), 595-618.

Knowles, J. B. *Doing better and feeling worse*. New York: Norton, 1977.

Koch-Weser, J. Validation of adverse drug reactions. In F. H. Gross and W. Inman (Eds.), *Drug Monitoring*, 1977, London Academic Press.

Kral, J. G., Lundholm, K., Bjorntorp, P., Sjostrom, L., and Schersten, T. Hepatic lipid metabolism in severe human obesity. *Metabolism*, 1977, *26*, 1025.

Krischer, J. P. Utility structure of a medical decision making problem. *Operations Research*, 1976, *24*(5).

Kunreuther, H. C., and Schoemaker, P. H. *Decision analysis for complex systems: Integrating descriptive and prescriptive components* (in press).

Laxdal, O. E., Jennett, P. A., Wilson, T. W., and Salisbury, G. M. Improving physician performance by continuing medical education. *Canadian Medical Association Journal*, 1978, *118*, 1051-1058.

Leon, G. R. Current directions in the treatment of obesity. *Psychological Bulletin*, 1976, *83*,(4), 557-578.

Lodwick, G. S. Computer-aided diagnosis in radiology: A research plan. *Investigative Radiology*, 1966, *1*, 72-80.

Maddox, G. L., Anderson, C. F., and Bogdonoff, M. D. Overweight as a problem of medical management in a public outpatient clinic. *American Journal of Medical Sciences*, 1966, *252*, 394-402.

Maddox, G. L., Back, K. W., and Liederman, V. R. Overweight as social deviancce and disability. *Journal of Health and Social Behavior*, 1968, *9*, 287-298.

Marschak, J. Actual versus consistent decision behavior. Behavioral Science, 1964, 9, 103-110.

McNeil, B. J., and Pauker, S. G. The patient's role in assessing the value of diagnostic tests. Radiology, 1979, 132(3), 605-610.

McNeil, B. J., Weichaelbaum, R., and Pauker, S. G. Fallacy of the five-year survival in lung cancer. New England Journal of Medicine, 1978, 299, 1397-1401.

Mikes, P. S., and Hulin, C. L. Use of importance as a weighting component of job satisfaction. Journal of Applied Psychology, 1968, 52, 394-398.

Neter, J., and Williams, C. A. Acceptability of three normative methods in insurance decision making. Journal of Risk and Insurance, 1971, 38(3), 385-408.

Newell, A., and Simon, H. A. Human problem solving. Englewood Cliffs, New Jersey: Prentice-Hall, 1972.

Nie, N. J., Hull, C. H., Jenkins, J. G., Steinbrenner, K., and Bent, K. H. Statistical Package for the Social Sciences. New York: McGraw-Hill, 1975.

Nisbett, R., and Ross, L. Human inference: Strategies and shortcomings of social judgment. New York: Prentice-Hall, 1980.

Pauker, S. G. Coronary artery surgery: The use of decision analysis. Annals of Internal Medicine, 1976, 85, 8-18.

Penick, S. B., Filion, R., Fox, S., and Stunkard, A. D. Behavior modification in the treatment of obesity. Psychosomatic Medicine, 1971, 33, 49-55.

Pliskin, J. S., Shepard, D. S., and Weinstein, M. C. Utility functions for life years and health status. Operations Research, 1980, 28(1), 206-224.

Price, P. B., Taylor, C. W., Nelson, D. E., Lewis, E. G., Laughmiller, G. C., Mathiesen, R., Murray, S. L., and Maxwell, J. G. Measurements and predictors of physician performance: Two decades of intermittently sustained research. Salt Lake City: L. L. R. Press, 1971.

Quereshi, M. Y. Some psychological factors that distinguish between the remediably and irremediably obese. Journal of Clinical Psychology, 1972, 28, 17-22.

Rabkin, S. W., Mathewson, A. L., and Hsu, P. H. Relation of body weight and development of ischemic heart disease in a cohort of young north American men after a 26-year observation period. The Manitoba Study. American Journal of Cardiology, 1977, 39, 452.

Ravitch, M. M., Rovner, D. R., Jennett, P. A., Rothert, M. L., Holmes, M. M., Holzman, G. B., and Elstein, A. S. A chart audit study of the referral of obese patients to endocrinologists (submitted for publication), 1982.

Rimm, A. A., Werner, L. H., von Yserloo, B., and Bernstein, R. A. Relationship of obesity and disease in 73,532 weight-conscious women. Public Health Report, 1975, 90, 44.

Rothert, M. L., Rovner, D. R., Elstein, A. S., Black, N. A., Jennett, P. A., Holmes, M. M., and Ravitch, M. M. Medical Referral Decisions as a Function of Professional Socialization. Paper presented at Annual Meeting of American Educational Research Association, March, 1982.

Schoemaker, P. H. Experiments on decisions under risk: The expected utility hypothesis. Boston: Martinus Nijhoff Publishing Co., 1980.

Shepard, R. N. Subjectively optimal selection among multi-attributed alternatives. In N. W. Shelley and G. L. Byron (Eds.), Human judgment and optimality. New York: Wiley & Sons, 1964.

Sisson, J. C., Schoemaker, E. G., and Ross, J. C. Clinical decision analysis. The hazards of using additional data. Journal of the American Medical Association, 1976, 236(11), 1259-1263.

Slovic, P., Fischhoff, B., and Lichtenstein, S. Response mode, framing, and information processing effects in risk assessment. In R. Hogarth (Ed.), New directions for methodology of social and behavioral science: Question framing and response consistency. San Francisco: Jossey-Bass, 1982, 21-36.

Slovic, P., and Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.

Smith, D. H., and McWhinney, I. R. Comparison of the diagnostic methods of family physicians and internists. Journal of Medical Education, 1975, 50, 265-270.

Solow, C., Silberfard, P. M., and Swift, K. Psychosocial effects of intestinal bypass surgery for severe obesity. New England Journal of Medicine, 1974, 290, 300-304.

Stimson, D. H. Utility measurement in public health decision making. Management Science, 1969, 16(2), B17-B30.

Stolley, P. D., Becker, M. H., Lasagna, L., McEvilla, J. D., and Sloane, L. M. The relationship between physician characteristics and prescribing appropriateness. Medical Care, 1972, 10(1), 17-28.

Taylor, T. R. Clinical decision analysis. Methods of Information in Medicine, 1976, 15(4), 216-224.

Torrance, G. W., Thomas, W. H., and Sackett, D. L. A utility maximization model for evaluation of health care programs. Health Services Research, 1972, 7(2), 118-133.

Tversky, A., and Kahneman, D. Judgment under uncertainty: Heuristics and biases. Science, 1974, 185, 1124-1131.

Van Itallie, T. B. Obesity: Adverse effects onn health and longevity. American Journal of Clinical Nutrition, 1979, 32, 2723-2733.

Von Neumann, J., and Morgenstern, O. Theory of games and economic behavior (3rd ed.). Princeton: Princeton University Press, 1953.

Von Winterfeldt, D., and Fischer, G. W. Multi-attribute utility theory: Models and assessment procedures. In D. Wendt and C. Vlek (Eds.), Utility, probability, and human decision making. Dordrecht, Holland: Reidel Publishing Co., 1975.

Vroom, V. H. Organizational choice: A study of pre- and post-decision processes. Organizational Behavior and Human Performance, 1966, 1, 212-225.

Weinstein, M. C., Fineberg, H. V., Elstein, A. S., Frazier, H. S., Neuhauser, D., Neutra, R. R., and McNeil, B. J. Clinical decision analysis. Philadelphia: W.B. Saunders Co., 1980.

Weinstein, M. C., Pliskin, J. S., and Statson, W. B. Coronary artery bypass surgery - decision and policy analysis. In J. P. Bunker, B. A. Barnes, and F. Mosteller (Eds.), Costs, risks and benefits of surgery. New York: Oxford University Press, 1977.

Weisinger, J. R., Kempson, R. L., Eldrige, F. L., and Swenson, R. S. The nephrotic syndrome: A complication of massive obesity. Annals of Internal Medicine, 1974, 81, 440.

Wendt, D., and Vlek, C. (Eds.) Utility, probability, and human decision making. Dordrecht, Holland: Reidel Publishing Co., 1975.

Williamson, J. W., Aronovitch, S., Simonson, L., Ramirez, C., and Kelly, C. Health accounting: An outcome-based system of quality assurance: Illustrative application to hypertension. Bulletin New York Academy of Medicine, 1975, 51(6), 727-738.

Winer, B. J. Statistical principles in experimental design. McGraw-Hill, 1971.

Yntema, D. B., and Torgerson, W. S. Man-computer cooperation in decisions requiring common sense. IRE Transactions on Human Factors in Electronics, 1961, 2(1), 20-26.