
UNIVERSITY OF NORTH TEXAS
HEALTH SCIENCE CENTER AT FORT WORTH
1970

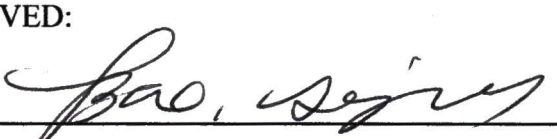
Whitesell, Rebecca, The Utility of Exploratory Data Analysis Techniques in Analyzing Outcome Measures Used in Osteopathic Manipulative Medicine Research. Master of Public Health (Clinical Research), December 2005, 79 pp., 2 tables, 42 illustrations, references, 33 titles.

The intent of this thesis project is to describe how Osteopathic Manipulative Medicine (OMM) researchers can use the theory of Exploratory Data Analysis (EDA) to enhance their ability to analyze research findings. This was achieved by evaluating the most frequently used outcome measures in OMM research published since 1993, describing EDA and its relevance to the types of data used in OMM research, and illustrating the ways EDA can be used in two current OMM studies to gain insights into the data and to shape future research questions.

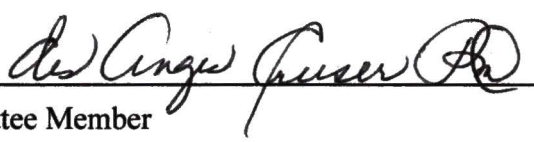
The Utility of Exploratory Data Analysis Techniques in Analyzing Outcome Measures Used in Osteopathic Manipulative Medicine Research

Rebecca Whitesell, B.S.

APPROVED:



Major Professor



Committee Member

Committee Member

Department Chair

Dean, School of Public Health

The Utility of Exploratory Data Analysis Techniques in Analyzing Outcome Measures Used in Osteopathic Manipulative Medicine Research

Thesis

Presented to the School of Public Health

University of North Texas

Health Science Center at Fort Worth

in Partial Fulfillment of the Requirements

for the Degree of

Master of Public Health

By

Rebecca Whitesell, B.S.

Fort Worth, Texas

December 2005

TABLE OF CONTENTS

Page

INTRODUCTION AND PURPOSE

The purpose of studying biostatistics as a major in a course of study leading to a Master's of Public Health is to ultimately apply the learned theories, statistical tools and techniques to actual research projects. This thesis project in biostatistics explains and describes how Exploratory Data Analysis (EDA) can provide a philosophy or mindset that enables a researcher to gain greater insights into data from a clinical trial that might not be readily apparent using traditional hypothesis testing methods. The focus of this project was on clinical studies of the efficacy of osteopathic manipulative medicine (OMM). The project was conducted at the University of North Texas Health Science Center (UNTHSC) that includes the School of Public Health (SPH), the Texas College of Osteopathic Medicine (TCOM) and the Graduate School of Biomedical Sciences. This project was conducted in TCOM because there is a national osteopathic research program housed in the Department of Osteopathic Manipulative Medicine. That research program, the Osteopathic Research Center (ORC) is at the forefront of research into the mechanisms of action and clinical efficacy of OMM, and OMM is one of the distinguishing features of Osteopathic Medicine.

The intent of this project was to produce a paper that could be used to teach OMM research fellows how EDA could enhance their ability to analyze research findings. This project was of interest to me because I have an undergraduate background in the basic sciences and have extensive course work in biostatistics, epidemiology and research design in the public health arena. It allowed me to combine my interest in medicine with my research and statistics background. I was particularly drawn to this project after

completing a practice experience at the ORC. During the practice experience I participated in data management and analysis for two active clinical trials.

The ORC is involved in over a dozen studies of the mechanism of action and efficacy of OMM. Their research teams include both clinician/physician scientists and basic (anatomy and physiology) scientists. To narrow the focus of this thesis project I participated in two of these studies to explore and describe the suspected utility of EDA in this type of research. Both trials in this project collected both objective and subjective data in their primary and secondary outcome measures. Because clinical trials in OMM tend to collect multiple types of outcome measures, I hypothesized that EDA would be the most useful method to help the research fellows 1) more thoroughly understand the shape and tendencies in their data, 2) improve the selection of hypothesis testing methods, and 3) generate better questions for future research.

There are two major aims for this project:

Aim 1: Describe the outcome measures that OMM researchers have primarily used over the past decade

Aim 2: Determine and describe the ways in which EDA can be used to gain important insight into mixed data sets, thus assisting the research in important ways

To achieve these aims I used four steps:

1) Identify and describe the most frequently used outcome measures described in OMM research published since 1993;
2) Evaluate outcomes used in the published research of the profession;

3) Describe EDA and its relevance to the types of data used in OMM research; and
4) Illustrate the ways EDA can be used in two current studies to gain insights into the data and shape future research questions.

BACKGROUND AND SIGNIFICANCE

This section discusses:

1. Definitions of project terminology and concepts
2. Principles and concepts underlying osteopathic medicine and OMM
3. Issues in outcome measurements in OMM mechanistic and clinical efficacy studies
4. Exploratory Data Analysis

Definitions of Project Terminology and Concepts

Outcome measures in clinical trials refers to any result that arises (in a patient) due to a specific intervention. This may range from improving physiological functions, such as blood pressure or forced expiratory volume, to more patient-centered outcomes such as pain levels, over-all well-being and quality of life (Kane, 1997). Other outcome measures may be length-of-stay, time to clinical stability, or reoccurrence of a condition or illness.

Exploratory Data Analysis is a statistical approach to data analysis that was developed by John Tukey in the 1960s and 70s. It, similarly to Osteopathic manipulative medicine, has been clouded by skepticism over its scientific rigor (Victor, 1982). EDA is not an independent statistical test, but rather a mindset with which one approaches data to

better understand its shape, tendencies, and underlying structures (Tukey, 1977; Hartwig & Dearing, 1979). The techniques of EDA are fairly simple (Shelly, 1996), rely heavily on visual inspection of the data, and can be used by researchers with only basic statistical knowledge. The foundations of the technique are in using visual displays of data to shed light on the relationships that exist among different variables (Tukey, 1977; Behrens, 1997).

Basic Science and Clinical Research

There are four important differences between basic science and clinical research studies that relate to this study. These are:

1. Clinical studies involve human subjects in a real-world (in vivo) setting, whereas basic science research is conducted under controlled laboratory conditions;
2. Clinical studies tend to use both objective and subjective measures of the hypothesized result or outcome of an intervention whereas basic science research relies primarily on objective measures;
3. Clinical studies have less ability to control the environment then basic science (bench) research;
4. Clinical studies often take the "bench" to "bedside" approach attempting to demonstrate that a particular theoretical or scientific phenomenon does have a particular desired effect on a health care population. This is known as "translational" research.

(1) Because clinical studies focus on the changes that may occur in a human subject following the administration of a treatment (Kane, 1997), the outcome measures

that will ultimately determine the answer can be both objective and subjective in the type of variables/data used. For an example, in cancer therapy research a scientist may hypothesize that a particular lab value will result from a particular treatment. However, the cancer investigator may also want to know whether the patient felt more or less pain, or was able to do more or less activities of daily living as a result of taking a new therapy. Both of these outcomes are important, and critical to a full understanding of the impact of the intervention.

(2) Lab values are typically ratio or interval. However, in measuring self-reported or clinician rated pain, for example, the researcher is less likely to use a standardized scale, and therefore has more variability in the data (Kane, 1997). Although pain can be measured somewhat objectively by some medical equipment, it is most often only measured indirectly because of the complex nature of the concept of pain. Generally studies of the impact of OMM are not easily measured objectively. This is a major challenge for OMM research for the future. Clinical outcome measures in studies of the efficacy of OMM have tended, in the past decade, to be more subjective than objective because very few objective outcome measures are available and validated.

The types of measurements used in subjective outcome instruments are difficult to analyze using traditional, hypothesis testing, statistical procedures, as they may violate the necessary assumptions. Exploratory Data Analysis provides a statistical tool that can be used to 1) visually analyze the data obtained from these studies, and 2) provide a context within which to generate more questions about plausible relationships. I do not suggest that these techniques be used as sole analysis tool in any study but I do suggest

that it is a useful approach to enhance the understanding of data obtained from research such as the preliminary, pilot-studies being conducted by fellows at the ORC.

Subjective measures are typically categorical variables with nominal and ordinal scaling, though some can be continuous variables, these are interval only if a range of scores is used such as in a Likert scale. Subjective measures address a variety of patient responses such as pain, satisfaction with treatment, severity of symptoms, level of daily functioning and degree of ease or difficulty with activities of daily living. Many researchers have traditionally discounted this type of measure because there is a great deal of variability among different patients. It is difficult to show, for instance, that two patients reporting the level of pain are actually feeling the same pain. Though Kane (1997) does state that subjective outcome measures are just as reliable as objective measures.

Objective measures traditionally are laboratory, fluid or machine obtained variables. These types of variables can be of any variable scaling type (nominal, ordinal, interval, ratio). Examples of objective measures include blood pressure, temperature, presence of disease state, and range of motion. These measures are referred to as objective measures because they exhibit some precise standardized quality or quantity in comparison to subjective measures that rely heavily on feeling and perception. Objective measures are based on external values, subjective measures are based on internal, individualized values.

Assumptions and violations of assumptions for parametric tests and non-parametric tests: Parametric tests are statistical procedures that test interval or ratio scaled

data. These tests assume that individual and group variables have a normal distribution with a mean of zero and equal group variance. When these two assumptions are violated, non-parametric tests must be utilized. Non-parametric tests are most commonly used for variables with nominal or ordinal scaled data. Parametric statistical procedures include independent and paired *t* tests, analysis of variance (ANOVA), and multivariate analysis of variance (MANOVA). Non-parametric tests include chi-square test, Mann-Whitney test, Sign test, Median test, Wilcoxon Sign Rank test, the Kruskal-Wallis test and the Kendall Coefficient of Concordance. One of the major advantages to using non-parametric tests is the absence of many of the restrictions required for parametric tests, mainly the normality and equal variance assumptions. The results obtained using non-parametric tests are, however, less specific in terms of what the results can describe about the data. (Neutens & Rubinson, 2002).

There are four types of measurement or scaling for data, variables or outcome measures called nominal, ordinal, interval, and ratio. Nominal data are divided into distinctly different categories such as gender or treatment group. Ordinal data is also divided into categories but the categories are arranged in order such that one category indicates a higher response than another. Examples of ordinal data include medication dosing level, level of education and illness severity. Interval data are continuous variables that are separated by how much they differ from each other. Height and weight are typical examples of interval data. Ratio data are the same as interval data but a true zero is included in the scaling of ratio variables. For instance, Celsius and Fahrenheit temperature scales are interval scales because zero degrees does not actually represent the

absence of temperature. The Kelvin temperature scale is considered a ratio scale because has a true zero value; zero degrees Kelvin is actually the absence of temperature where this is not so in the other temperature scales (Neutens & Rubinson, 2002). Nominal and ordinal data are often referred to as categorical variables and interval and ratio data are called continuous variables (Hulley et al., 2001).

The Principles and Concepts Underlying Osteopathic Medicine

Osteopathic Medicine has a long history of struggle to gain respect in the scientific community. Since the days of Andrew Taylor Still, M.D., Osteopathic Medicine's founding father, and his endeavors to develop the first Osteopathic schools of medicine as well as his debates with the American Medical Association, this branch or medical science has been working in many venues to validate its principles. The story goes that A.T. Still was first disillusioned by allopathic medicine's philosophical underpinnings when he witnessed fellow physicians' use of toxic medications in the mid-19th century. Still was stupefied by the use of untested, unregulated, and often lethal doses of medication. He felt there must be an alternative to treating patients that was not solely focused on the elimination of disease through medication but also concentrated on the overall health of the individual (Jones, 1978).

Still developed a school of medical thought that treats patients with the philosophy that aligning the body to its correct, natural physiologic state would defend the body against disease and other aliments. This physiologic correctness was discovered by Still when, as a young boy, he relieved himself of a headache by resting his neck in the rope of a tree swing. Still used this experience as a foundation to develop a series of

manual procedures performed by physicians that are now called Osteopathic Manipulative Treatments. The physician is able to manually examine a human body for evidence of tissue changes and precisely apply procedures developed by Still to align the body to a state of physiologic correctness (Jones, 1978). Texts in OMM explain the theoretical basis for how OMM makes changes in the body's musculoskeletal and fluid/physiology systems (DiGiovanna & Schiowitz, 1991). Research is limited demonstrating actual effects theorized in the texts.

Issues in Outcome Measures in OMM Clinical Research

Because of the principles associated with "wellness" and the "body's natural ability to heal" and "musculoskeletal alignment – or lymph flow" (DiGiovanna & Schiowitz, 1991) it continues to be a challenge to test the theories underlying the mechanisms that OMM physician believe are actually affected by OMT. Furthermore, OMM is not normally a solitary treatment for disease; it is adjunctive, or in addition to other medical treatments making it more difficult to associate an outcome directly with OMT (Jones, 1978). Musculoskeletal problems or conditions such as low back pain or neck pain are among the more obvious candidates for OMM. Diabetes, neuropathies, sinus problems, gastrointestinal, or cardiovascular problems are less often associated with a potential benefit from OMM, but are in fact among the disease processes being studied in OMM efficacy studies today (Johnson & Kurtz, 2002). Depending on the condition that is targeted by OMM, various outcomes may be important to the researcher.

Much of the concern in regard to OMM research is related to the use of subjective versus objective outcome measures. Some of the most commonly used outcome

measures in OMM are subjective measures, which are often criticized because they rely on patient self-reported information. There is a scientific and professional bias that believes these measures to be inferior, in regard to reliability and validity, compared to objective measures using standardized values (Kane, 1997). Subjective measures differ from objective measures in that they use nominal, ordinal or dichotomous variables rather than ratio or interval scales. Statistically the use of subjective measures limits the investigator to a smaller number of possible statistical analyses that are often nonparametric in nature (Neutens & Rubinson, 2002).

In order to improve the quality of research performed in the field in a short time with many pressures from the profession and no previous track record with the NIH, the ORC strives for rigor in the design of OMM clinical trials and drives toward publication of reliable findings. Research endeavors that have taken precedence in terms of funding and support are prospective, randomized, blinded, controlled trials (PRBC) that study the aspects of OMM that have a preponderance of evidence that suggest or justify more advanced study. This includes researching the role of OMT in OB, CABG, sympathetic nervous system, carpal tunnel syndrome, pulmonary system, pneumonia, otitis media, and low back pain. Though the ORC has increased awareness about clinical research in the OMM profession there are still only a small number of studies being conducted each year.

Exploratory Data Analysis

The concept and techniques of EDA were first described by John Tukey in the 1960s and 70s. EDA is a set of tools that can be used by researchers to conduct data

analysis and it is also a philosophical mindset or attitude in terms of how one thinks about the data and conducts the analysis (Tukey, 1977; Behrens, 1997; Lederman, 1992). Lederman (1992) suggests that EDA focuses on visually analyzing the data either in the absence of or in conjunction with confirmatory statistical analysis. Behrens (1997) characterizes EDA as emphasizing and understanding what is "happening" in the data, graphically representing the data, focusing on model building and hypothesis generation, using robust measures and maintaining a position of skepticism and flexibility regarding the methods to apply.

EDA can be considered a hypothesis generating form of data analysis; hypotheses can be created from what is learned by EDA but the technique cannot be used to either accept or reject a null hypothesis, as in confirmatory data analysis (Behrens, 1997). Ferketich and Verran (1986) express a similar conceptualization of EDA in that it can result in theory generation "or theory testing can be enhanced by planning the appropriate confirmatory statistic for testing the developed hypothesis (p. 465)." In general, EDA is a tool to be used by researchers to gain a more comprehensive understanding of individual variables, relationships between variables and relationships that exist across all data.

Tukey (1977) proposed that researchers should approach and examine data with skepticism and open mindedness, rather than narrowly focusing on only testing the hypothesis. Tukey likened the methods of EDA to detective work. This is different from what has pejoratively been referred to as a "fishing expedition". Lederman (1992) suggests that researchers who remain doubtful and unconvinced by visual and statistical

summaries and question the accuracy of the original data exhibit these traits of skepticism and open mindedness. Measures that summarize the data, such as simple descriptive statistics such as measures of central tendency, should also be viewed skeptically because they can easily hide and misrepresent important features of the data (Ferketich & Verran, 1986).

The idea of openness when using EDA pertains to the researcher's ability to remain open to any unexpected patterns or relationships in the data. The necessity of the investigator to remain open and skeptical while utilizing EDA techniques is very important because, as suggested by both Ferketich and Verran (1986) and Lederman (1992), EDA analyses may reveal interesting aspects of the data. Ferketich and Verran (1986) further exemplified this idea when stating "The purpose of EDA is to obtain maximum information about the unexpected and to organize the data into meaningful patterns (p. 465)."

Since Tukey's original discussion on EDA, there have been many researchers who have summarized the techniques in general terms for other researchers in specific fields of study, such as nursing and psychology. I have also attempted to create a summary of EDA for OMM clinical research by both using Tukey's text as well as the summaries written by other researchers. In my opinion, Hartwig and Dearing (1979) and Behrens (1997) have written the most thorough descriptions of Tukey's exploratory methods. As Behrens (1997) suggests "the use of exploratory procedures such as plotting of simple summaries or the tabulation of simple descriptive statistics does not necessarily imply EDA (p. 134)." Meaning that just because a researcher follows the techniques

described does not imply that one has completed EDA; the philosophy is what researchers should adhere to because the specific techniques of EDA are secondary to gaining a rich description of the data (Behrens, 1997).

In subsequent sections of this paper (Findings) the techniques of EDA most commonly described in the literature will be elucidated using examples from research conducted at the ORC. The techniques used in these examples are not the gold standard for EDA because there is no gold standard, though they do represent the techniques commonly used in EDA as identified in the literature. The goal of EDA is to help the researcher understand patterns in the data. Thus the visual analyses that provide the individual researcher with the best possible understanding of the data are the tools that should be used by that particular researcher. Given the same data, two separate researchers may utilize different visual displays to examine the data, though both researchers could still be utilizing EDA principles (Behrens, 1997). Essentially, as long as the researcher examines the data with some degree of doubt and open mindedness there is no wrong way to use EDA. The only way in which EDA can be used inappropriately is if it is used at the conclusion of data analysis, after significant findings have not been found, as a way to find any relationship among the data even if it does not necessarily pertain to the study (the pejorative "fishing expedition") (Victor, 1982).

METHODS

When I first began my practice experience and started to help the fellows with the analysis of their data, I was very concerned and had a great deal of apprehension about

my abilities to help medical students better understand their data. Despite my acceptance into medical school, I do not at this time in my training have much exposure to actual clinical trials. My experience has been in studying the theory of statistics not its practical use. EDA was a new concept to me and I was uncertain that it would be useful to the analysis of clinical data. I focused on using my biostatistics knowledge and tried to provide practical applications of the information I had learned in class to their research questions. However, I began to appreciate and value the utility of EDA because I had first hand experience with unexplainable findings using traditional hypothesis-testing methods in the available study data.

For this project I used a four-step approach to understand the two clinical trials in order to better perform applicable statistical analyses. These steps were to:

1) Collect information about the design of the study
2) Learn what outcomes were being measured
3) Conduct literature research into the types of analyses that were possible
4) Apply the techniques of EDA to each study

The first step was achieved by meeting with the investigators of the two studies to discuss the research studies, hypotheses and basic design of the study. I also read the research proposals for both the studies and determined the outcome measures recorded for each study.

For this project I reviewed all published OMM clinical trials for the past 10 years to identify the outcome measures most commonly used in OMM research and EDA methods illustrated by focusing on two current clinical trials conducted at the ORC.

Literature searches were conducted in the three main areas concerning this paper: EDA, OMM PRBC trials and outcome measures utilized in OMT studies. The terms randomized controlled trial, OMT and outcome measures were entered, in both Ovid and OstMed databases which resulted in a small number of applicable results in regards to publications concerned with PRBCs studying OMT. Seemingly, at least in relation to the number of publications, the most extensively studied application of OMT is in the treatment of low back pain.

FINDINGS

This section of the paper is divided into two major sections. The first discusses the findings from the research into the principles of basic and clinical research as well as a literature search focused on both outcome measures in OMT clinical trials and EDA techniques. The second section of this paper focuses on the results from the application of EDA techniques to the two current clinical trials conducted at the ORC.

Literature Research Findings

Basic Science versus Clinical Science Research

When designing research studies Punch (2000) outlines four main steps in defining the research model. These steps are as follows:

1) Conceptualize the research in accordance with the research questions;
2) Determine the necessary data needed to answer the research questions;

3) Design the research methods that will allow for the collection and analysis of the data; and

4) Answer the research questions based on the findings from the data analysis.

(1) Clearly defining the outcomes to be investigated in a study and establishing the form in which these variables will be measured (nominal, ordinal, interval or ratio scaling) is extremely important at the onset of the trial. Identifying the variables of interest by stating primary and secondary questions will focus the researcher's data collection process and make analysis, at the conclusion of the trial more complete. Defining the questions to be answered along with the variables and their subsequent scale type will aid in conducting the correct type of power and sample size calculations. If a researcher does not conceptualize the research design through the final stages of developing the data analysis plan and understanding its implications on power and sample size, any weaknesses in the study design will most likely not be discovered until the study has concluded and the collected data is analyzed. In fact it is unfortunate that such weaknesses occur in study designs because they are preventable (Punch, 2000).

(2) The data management plan is what drives much of the study. Two disadvantages that may occur as a result of a weak study design include 1) power and sample size inaccuracies and 2) incomparable variables during data analysis. Power and sample size inaccuracies are illustrated in the following example. If the researcher initially states the data will be collected in preparation for survival analysis and the power and sample size calculations are based on this type of analysis but in actuality a different type of analysis is performed, the sample size and power may be either inadequate or

overly-adequate for the new type of analysis. Underestimating or overestimating the correct sample size needed for a trial are both inappropriate because it may place subjects at risk and waste resources. Subjects are placed at risk in either situation because underestimation, of sample size, makes any conclusions from the study non-generalizable and the results are not credible in regard to the validity of the findings due to inadequate sample size. An overestimation also places subjects at risk due to possible adverse reactions to the treatment. Improper design of the study also wastes vital resources including funding, personnel and depletes the pool of potential subjects for other research investigations (Hulley et al., 2001; Neutens & Rubinson, 2002).

The second disadvantage that may arise from inadequately developing the data analysis plan during the research design phase includes incomparable variables. Because different variable scaling types require different types of analysis procedures, such as the use of parametric versus nonparametric statistical procedures, not determining the scaling of variables prior to conducting the trial may make comparison of certain variables statistically impossible. The comparison of a nominal variable and a ratio variable may be impossible without reorganizing the ratio variable into a comparable form such as nominal or ordinal. Though more descriptive variables such as interval and ratio variables can be reorganized by the researcher into nominal and ordinal variables particularly vital information may be lost by confining a continuous variable to categories (Neutens & Rubinson, 2002).

Basic science researchers have set a standard that is not necessarily suitable for all clinical trials. As a result, clinical researchers may conduct "hypothesis testing" rather

than a thorough, meaningful examination of results. It is possible to examine results and find clinical importance without having to prove an event that occurs by chance will occur in only 5% of the cases. In some situations 10 percent or 12 percent of such a chance event may be adequate. As in any human process, a 90 percent chance of a positive result, is an overwhelming probability that something good will occur, especially in the public health arena, is often enough to put the action into (Abelson, 1995; Denis, 2003; Froehlich, 2004). A possible explanation to clinical researchers' zealous attempts to obtain results with *p*-values below 0.05 is not only due to influence of bench science but some have also suggested a publication bias toward research that meets this level of significance. Just because a treatment option, technique or procedure is show to be statistically insignificant does not indicate that it may not be clinically important. Numerous authors have described the problems associated with using *p*-values to measure clinical importance (Abelson, 1995; Denis, 2003; Froehlich, 2004) and Froehlich (2004) has even suggested a new test for clinical importance called *q*-values.

Outcome Measures Used in OMT Clinical Trials

I was able to identify ten published PRBC trials examining the efficacy of OMT in treating various conditions published since 1993. The majority of the identified trials were published in the Journal of the American Osteopathic Association (JAOA), and others were found in Spine and the New England Journal of Medicine. Some of the studies published in JAOA were full articles (two) while others (six total) were only published poster abstracts from the annual AOA conferences.

There were two randomized studies of efficacy of OMT in low back pain, one by Andersson, et al. (1999) and one by Licciardone et al. (2003). There was also a trial concerned with OMT and hospitalized elderly pneumonia patients by Noll, et al. (2000) and a study of OMT and fibromyalgia syndrome by Gamber et al. (2002). Poster abstracts from the annual AOA meetings published in JAOA, included studies examining the efficacy of OMT in recovery time from surgery, pneumonia, relief of headache pain, Parkinson's disease, childhood otitis media, and hip surgery. The discussion of these trials is limited to the outcome measures utilized by each. A table summarizing the measures used in each trial is shown as Table 1 in the Appendix.

Outcome Measures Identified in Published OMT Studies

Outcome measures used in the two low back pain studies are both similar and dissimilar. "Osteopathic Manipulative Treatment for Chronic Low Back Pain: A Randomized Control Trial" (Licciardone et al., 2003) used fewer reported outcome measures than the Andersson et al. (1999) study ("A Comparison of Osteopathic Spin Manipulation with Standard Care for Patients with Low Back Pain.") Licciardone (2003) recorded standard demographic information and five descriptive or classification variables. Outcome measures used in this study include health status, two measures of pain and disability as well as number of lost work or school days within a four-week time period and patient satisfaction with treatment.

The Andersson (1999) study used six classification variables and seven outcome measures including pain and limitations/disability, "acceptance" of pain, pain drawings, range of motion and straight leg raise. Both of these studies examining the efficacy of

OMT in low back pain utilized subjective outcome measures. Andersson (1999) included two objective outcome measures: range of motion and angle of degree of a straight leg raise.

"Benefits of Osteopathic Manipulative Treatment for Hospitalized Elderly Patients with Pneumonia," by Noll et al. (2000) used both objective and subjective outcome measures: demographic information, where pneumonia was acquired (community, nursing home, or hospital), severity of illness and patient vital signs. Severity of illness was estimated by the Simplified Acute Physiology Score (SAPS) and patient vital signs including temperature, pulse, and respiratory rate were recorded and periodic x-rays were studied for evidence of pulmonary infiltrates.

The Gamber (2002) study ("Osteopathic Manipulative Treatment in Conjunction with Medication Relieves Pain Associated with Fibromyalgia Syndrome: Results of a Randomized Clinical Pilot Project") assessed three measures of pain (ten tender points, Chronic Pain Experience Inventory and Present Pain Intensity Rating Scale). Response to treatment, activities of daily living and depression were also recorded using a Self-Evaluation Questionnaire, Stanford Arthritis Center Disability and Discomfort Scales: Health Assessment Questionnaire and the Center for Epidemiological Studies Depression Scale, respectively.

Outcome Measures Identified in Published OMT Research Abstracts

The studies reported in abstracts also have various and mixed outcome measures. A study called "The Effectiveness of Osteopathic Manipulative Treatment as Complementary Therapy Following Surgery: A Prospective, Match-controlled Outcome

Study," (Jarski, 2000) included outcome measures of days to independent negotiation of stairs, distance ambulated, supplemental intramuscular analgesic use, length of hospital stay and patients' perceptions of treatment. The abstract by Noll, et al. (1998), "The Efficacy of Adjunctive OMT in the Elderly Hospitalized with Pneumonia" included outcomes of duration of IV and oral antibiotic use as well as the length of hospital stay. In the headache pain relief study entitled "Evaluation of the Effect of Osteopathic Manipulative Treatment on Headache Pain: Duration of Pain Relief" conducted by Handler et al. (1998, p. 390) used "survey forms" to assess pain and symptoms, and medication use following treatment, as the study outcomes.

The study of OMT in Parkinson's disease patients entitled "Osteopathic Considerations in Parkinson's Disease," by Smutny et al. (1998) evaluated three dimensional gait analysis, general fitness, cardiovascular fitness, general health, osteopathic structural analysis, nutrition assessment and assessment of home environment. The study conducted by Brittain et al. (1997) on the efficacy of OMT in post-hip and knee surgery at a rehabilitation center used a standard disability measurement tool and the Functional Independence Measure, discharge destination and length-of-stay as clinical outcome measures.

The published poster abstract of OMT and otitis media (Steele, et al., 1997) named "Effect of Osteopathic Manipulative Treatment on Childhood Otitis Media Outcomes" utilized five outcome measures. 1) An eight-question scale was administered at each visit to address the subject's behaviors known to be associated with otitis media. 2) Tympanometric measurements which included measures of adequacy, baseline

adequacy peak, VEA, TPP and classification, 3) the number of office visits for otitis media, and 4) the number of prescriptions dispensed for the illness and 5) surgical referrals were recorded as outcome measures for each patient.

Exploratory Data Analysis

This section discusses the literature findings concerning EDA. Most authors have divided the techniques of EDA into two applications, 1) those used to examine single variables and 2) those to compare variables against each other. This discussion will also be divided into these two topics.

Examination of Single Variables

Most authors have suggested that single variables be examined using various displays to show the distribution of the variable under study, including histograms, box plots and stem-and-leaf displays. These visual displays are used to characterize the variable's location, spread and shape, and to search for the presence of any outliers. The *location* of the variable is determined by identifying the central point in the distribution. The *spread* pertains to how the distribution is dispersed and the width of such dispersion, i.e. is the variable distributed within a narrow area of possible values or spread over a wide range of possible values. The *shape* of the distribution refers to how the data points are configured; for example they could be distributed along a straight line, in a curve, which could be bell-shaped with one peak or multiple peaks, skewed or rectangular. Outliers are those values in the distribution that are extreme values compared to the other values obtained. They could be either much more or much less than other obtained or expected values (Lederman, 1993; Verran & Ferketich, 1987a,b).

Methods to Analyze Single Variables

There are three methods that are most commonly identified to visually examine the distribution of single variables. The three most commonly used are the box plot, stem-and-leaf display and the quantile plot. Other measures of distribution include the histogram, median hinge number summary and symmetry plots.

The stem-and-leaf display, according to Verran and Ferketich (1987b), is very useful when ordering the data and visually representing the shape and distribution of the variable. This type of display is created when the numbers are placed on the left side of the plot, these are called the stems and are intervals found in the data. The leaves are on the right side of the graph and represent the individual value of each data point within each interval. A stem-and-leaf display looks similar to a histogram but the exact values of each observation are maintained where a histogram shows the frequency of data at each interval.

Verran and Ferketich (1987b) liken the quantile plot to the stem-and-leaf as a good initial way to view the distribution, and more clearly identify the median and 25^{th}, 50^{th} and 75^{th} percentiles. "A quantile is a fraction of the data that represents the score's location in the distribution and may be calculate from the formula $(i\text{-}0.5)/n$ where i is the order of the score and n is the number of subjects (Verran & Ferketich, 1987b, p. 143-144)."

The box plot is an excellent way to look for outliers (Shelly, 1996). This method displays the 75^{th} and 25^{th} percentiles (interquartile) as the outer edges of a box with the median of the distribution dividing the box in two. Tails extend from the box,

traditionally called whiskers, which reach to the edge of the distribution represented by 1.5 times the interquartile range value (distance between the 25^{th} and 75^{th} percentiles). Any values that extend beyond the whiskers are termed outliers and are labeled by symbols (commonly asterisks) in the display.

The histogram, median hinge number summary and symmetry plots are also techniques available to examine single variables. The histogram resembles the stem-and-leaf diagram in shape and distribution but the individual characteristics of the values within each interval are not maintained. Each bar in the histogram represents the number of points in the distribution that fall within each specific interval. This type of plot aids in viewing the distribution and shape of the distribution of the variable (Shelly, 1996).

The median hinge summary is, in my opinion, the least visually understandable of the methods described in the literature. The diagram is comprised of three rows; the first row shows the lower extreme, lower hinge, median, upper hinge and upper extreme values. The second row represents the spreads of each quartile, and the third row shows the spread between the lower extreme and the median, the lower hinge and the upper hinge and the median and upper hinges (Verran & Ferketich, 1987b).

A symmetry plot compares two values in the distribution against one another. The two most extreme values are plotted against one another, then the second two most extreme values are plotted against each other and so on. The final plot should be a straight line if the distribution is symmetrical (Verran & Ferketich, 1987b).

Examination and Methods to Compare Variables

The relationship between variables has three important characteristics that include shape, strength and direction (Lederman, 1993; Verran & Ferketich, 1987a). The *shape* of the distribution is described the same as for a single variable. The *strength* is the correlation between the two variables and the *direction* refers to how the two variables are related, such as if the high values of one variable are related to the low values in the other variable. According to most authors the best way to visually look at the relationship between two variables is with a scatterplot, which is simply a plot of all the values within each variable on an x-y axis (Shelly, 1996; Verran & Ferketich, 1987a).

Additional Characteristics of EDA

Overall, the specific techniques employed by EDA are relatively simple, from a statistical standpoint, and require little formal statistical background (Shelly, 1996). Though it may be important to understand the concepts of variables, some geometry, and some graphic representation of numbers as well as the scaling of each variable. Though these techniques are relatively easy for most researchers to understand in comparison to formal hypothesis-testing procedures, and are available and applicable to all research, the use of EDA is not widespread. Similar to the controversy that exists among those in the Osteopathic profession regarding the efficacy of OMT, many statisticians do not view EDA as a valid method of analysis (Victor, 1982). The most common objection suggests that if one looks at a particular set of data enough a potentially significant relationship will eventually appear (Shelly, 1996). Ferketich and Verran (1986) noted this in stating, "...the importance of visual displays and resistant statistics is often degraded and

considered less scientific and less objective than other methods. Although it is true that there is an element of subjective interpretation to graphical analysis, this is minimized when a series of visual displays continues to evidence the same pattern of clues to the data (p. 465)."

A major problem in which those who advocate the use of EDA must combat are those researchers who suggest they have used EDA but in fact are only using exploratory methods like simple outlines of descriptive statistics and plots. Using the methods of EDA does not imply the use of EDA. Victor (1982) makes two suggestions on how to "prevent the abuse that would damage the reputation and spread of EDA (p.54)." First, Victor (1982) states that experiments must be thoughtfully planned with questions to be asked, response variables, and populations identified before the experiment is conducted. Second, the discussion of results should include all analyses so not as to obscure a particular finding.

A major limitation to EDA and possibly one of the reasons this approach to preliminary analysis is not used by more investigators may be due to the fact that examining the data with EDA alone proves nothing. Inherent in the philosophy of EDA is that it only generates hypotheses and theories, it cannot prove them; confirmatory data analysis must be used to prove or disprove any relationship among the variables (Behrens, 1997). Authors such as Victor (1982) suggest that this is just the value of EDA because without its methods the confirmatory statistical test could never generate new hypotheses or identify a new model.

Applicability of EDA to OMM Research

The methods of EDA are applicable to research projects in OMM. Many researchers feel it necessary to conduct hypothesis-testing, conformational analysis techniques for all clinical research even when it may not be applicable. Behrens (1997) addresses this dilemma when suggesting, "mathematics should be used based on how helpful it is in understanding data, not simply on its syntactical correctness (p. 155)." The problem of indiscriminately applying hypothesis-testing statistics is particularly harmful in preliminary and pilot studies that focus on trying to understand relationships among variables and generate more hypotheses. Often, when traditional hypothesis-testing procedures are applied to these studies and the tests and relationships are shown to be statistically insignificant, little to no subsequent research is performed on this topic. Verran and Ferketich (1987a) propose that placing an emphasis on EDA "can enhance a program of research in ways that could not ordinarily be available if only nonresistant numeric summaries were used (p. 623)." Thus by encouraging the use of EDA among investigators searching to explain the potentially beneficial role OMT could play in the treatment of various diseases, it is possible that the data could be better understood, more hypotheses could be generated and more advanced research could be conducted and tested using confirmatory, hypothesis-testing statistics.

Victor (1982) believes that since clinical research is so imprecise and no single hypothesis could ever by generated, using EDA is "an indispensable and useful aid for the discovery of new phenomena (p. 54)." Other fields performing clinical research such as nursing rely on exploratory techniques to tap into the vast potential of information

obtained from patient research. Since there have been no studies looking at the applicability of EDA to OMT research, the understanding of the utility of EDA in nursing techniques research is important. Both Lederman (1992) and Ferketich and Verran (1986) made reference to its applicability in the field of nursing. Lederman (1992) suggests the field of nursing is an evolving science that is continuously re-evaluating and defining its research domain and that EDA is the perfect tool to learn as much as possible from data so new theory can be developed, tested and refined. Ferketich and Verran (1986) state, "EDA provides a unique tool for nursing research. Human responses do not always follow regular patterns and may have many deviations from the expected. The location and analysis of this randomness is a valuable contribution to science and may help explain the unknown portion of the universe of interest (p. 466)."

"A powerful strategy of data analysis is ignored if EDA is not considered in the research program (Ferketich and Verran, 1986, p. 466)." They also suggest that employing traditional confirmatory statistics to a rich patient database neglects the potential of finding relationships among the data. This is particularly true especially in preliminary studies in which one of the main objectives in conducting the research is to generate more questions. Hypothesis-testing statistics only allow the acceptance or rejection of the null hypothesis. By using EDA techniques in conjunction with hypothesis-testing these rich databases may be utilized to their full potential if researchers investigate the relationships that may exist between any number of variables in the data.

If EDA methods were to be used in OMM research it is possible that the data could be more thoroughly examined by investigators, and new hypotheses and theories

could be generated. Ferketich and Verran (1986) also suggest that patient self-reported measures do not follow predictable patterns. There is some beauty in using subjective measures because they allow us to discover the true fullness of variability in the responses provided by different people. The strategies of EDA are helpful in understanding what makes self reports so variable, as well as how this potential variability may influence other measures.

There seems to be little preventing researchers from using EDA in conjunction with traditional confirmatory hypothesis testing procedures. The additional time it may take researchers to perform EDA is minimal in comparison to the benefits, including being able to approach the data in a more organized way.

Results of Applying EDA to Current Study Data

The aims of this project are to provide an understanding of outcome measures in OMT, the principles of EDA and how EDA can be used in OMT research. The final material in this section is devoted to a description of how EDA was used in two research projects currently being conducted at the ORC. To illustrate the aforementioned EDA strategies, each of the techniques were applied to the research data from the Carpal Tunnel Syndrome and OB studies. Graphs and other visual data displays for both studies are discussed. All visual displays are presented in the Appendix.

OMM and Carpal Tunnel Syndrome

This study was designed to test the efficacy of OMT in Carpal Tunnel Syndrome (CTS). The study has two main parts. The first part of the CTS study is concerned with the effects OMT has on nerve conduction and edema in the carpal tunnel. The second

part of the study is concerned with understanding how OMT may produce changes in symptom severity, functional status and strength. Part one has two hypotheses, 1) "OMT will improve the electrophysiologic conduction of the median nerve over a subacute interval as measured by nerve conduction studies (Meyer, 2004, p. 15)." 2) "OMT will produce subacute changes in the cross-sectional area of the carpal canal and fluid content (edema) of the median nerve as measured by magnetic resonance imaging (Meyer, 2004, p. 15)." The outcome measures used in the first part of this study are objective measures recorded from tests monitoring nerve conduction and MRI reports. Part two of the CTS study also has two specific hypotheses, 1) "OMT will decrease pain and other symptoms of CTS patients measured by the visual analog pain scale and the Carpal Tunnel Symptom Severity Questionnaire (Meyer, 2004, p. 15)." 2) "OMT will improve the daily functioning level of subjects with CTS as measured by the Functional Status Questionnaire and grip and pinch strength (Meyer, 2004, p.15)." This part of the study used the following outcomes: patients' self-reports of symptom severity and functional status and mechanically derived indicators of strength.

The CTS study called for 50 subjects, one-half in the ultrasound, placebo-control group and half in the OMT treatment group. The research staff and patients are blinded to treatment group but the physicians and student physicians providing treatment know the group assignment. This unblindedness of the physicians and student-physicians is necessary in order for treatment to be provided to the subjects. The treatment group received OMT "to the general areas of the wrist, arm, shoulder, neck and back in addition to any current standard care as outlined by the subject's primary care physician (Meyer,

2004, p. 3)." The control group received a placebo treatment that consisted of sub-therapeutic ultrasound to the same general areas in which the treatment group received OMT.

OMM and Pregnancy

The second study (Licciardone, 2004) considered in this paper is concerned with the efficacy of OMT on low back pain experienced during pregnancy. This trial is a randomized, placebo-controlled, blinded trial. There are three possible treatment groups, one receiving OMT, one receiving sub-therapeutic ultrasound, and the third receiving standard OB care. The OMT treatment and sub-therapeutic ultrasound groups will receive treatment to the neck, back, arms, legs and pelvis. The outcome measures used in this study include the Roland-Morris, the Quadruple VAS, SF-12v2 Health Survey, and confidence in treatment questionnaires. The hypothesis of this study is that subjects receiving OMT will experience less low back pain during pregnancy, delivery and post partum and have other better health indicators during delivery.

Analysis of the study data using EDA

Exploratory Data Analysis was used to analyze the data from both studies available at the time this paper was written. Both studies had enrolled nearly the same number of patients (n=25); this constituted approximately 50% of the projected enrollment for the CTS study and 25% for the OB study. Due to the large number of variables utilized in the CTS and OB studies a table of all the variables used in these two studies is provided in the Appendix of this paper.

Because the goal of this paper is to familiarize researchers with the potential applications of EDA in studies of OMM, and other research that incorporates subjective outcome measures, I have chosen to show how EDA procedures were used in the CTS study. The variables selected for illustration of EDA techniques include strength measures (grip strength), symptom severity scores, and visual analog scores. The strength measures represent objective, mechanically derived, the symptom severity and visual analog scores are subjective outcome measures.

Visual analysis procedures followed those described under the explanation of EDA for both the investigation of single variables and for examining two variables together under the subheadings Methods to Analyze Single Variables and Examination and Methods to Compare Variables. All EDA procedures were performed using SPSS Grad Pack 10. Box plots, stem-and-lead displays, histograms, bar graphs and scatterplots were produced for each outcome measure. Quantile plots, median hinge number summaries and symmetry plots were also discussed in the EDA section but these summary methods were unavailable using SPSS. The absence of these plots seems to be acceptable when reviewing examples found in the literature because these types of plots seemly provide less visual understanding of the data than those obtained using SPSS and many aspects of the variable distribution these plots could show are expressed in the other visual displays.

Bar graphs were not described in the literature but they are a very simple way to represent data. These types of graphs can be organized in various ways, showing single variables or variables clustered together. Numerous bar graphs have been included. In

my opinion they seem to be the most recognizable and easiest to understand by individuals of various backgrounds; they are a very clear and easy way to represent data as single variables or multiple variables. These types of visual analysis seem to bridge the gap between those techniques used to look at single variables and those used to look at multiple variables.

CTS

A) Strength Measures

The first series of displays to discuss are those produced for the strength measures; grip strength was used for explanatory proposes of this paper. During the study the practitioner recorded these strength measures at three of the patient visits (visits 3, 6 and 9). Visits 3,6 and 9 are referred to as grip strength 1,2 and 3 respectively, in the various displays and subsequent discussion.

Figure 1a is a box plot representing the three grip strength measures. The box visually surrounds the middle 50 percent of the measured points (the bottom and the top represent the 25^{th} and 75^{th} percentiles), this is also known as the interquartile range. The line dividing the box represents the median of each grip strength measure. The tails or whiskers extending from the boxes represent the points that fall within 1.5 times the interquartile range. Outliers are noted by individual markers, this plot has no outliers. There seems to be a progressively, gradual increase in the median grip strength value between grip strength 1 and grip strength 3. A question that may arise when looking at this plot in terms of the research question pertains to whether there is a difference in grip strength scores between treatment groups.

Figure 1b is a box plot that represents each of the three grip strengths based on treatment group (OMT or ultrasound) and aims to answer the previously stated question regarding grip strength scores and treatment groups. This group of box plots suggests that the median grip strength gradually increases from measurement 1 to measurement 3 in the OMT group. The median grip strength in the ultrasound treatment group increases from time 1 to time 2 and then decreases in time 3, to a value similar to that observed in time 1. This suggests that OMT may be providing positive benefit in terms of grip strength. This observation, in itself, obviously does not provide the investigator enough information to draw a conclusion about the efficacy of OMT treatment. More visual analyses will help to provide a more thorough understanding of the variable as well as generate more questions.

Stem-and-leaf diagrams for grip strength 1, 2 and 3 are provided in figures 2a and 2b. Those in figure 2a represent the average of grip strengths 1, 2 and 3. Figure 2a has each grip strength divided by groups and separate stem-and-leaf diagrams are produced for each strength measure recording (1,2 and 3) under each treatment regime. The stem-and-leaf diagram displays each data point (leaf) as an extension of its base value (stem) to better visualize the distribution of the data points.

Histograms display the same information presented in the stem-and-leaf plots but they do not provide the value of each individual data point in the visual presentation. Histograms for the average grip of strengths 1, 2 and 3 are found in figure 3a and histograms separated based on treatment group are seen in figure 3b. These histograms were produced with a superimposed normal curve to help visually indicate skewed data.

Figure 3a shows that the average grip strength has an approximately normal distribution, with scores that have a range of approximately 45 points between the highest and lowest score. The histograms contained in figure 3b show each grip strength test 1,2 and 3 based on treatment groups, both measures have an approximately normal distribution. The means generally increase from measure 1 to measure 3 in the OMT group but stay approximately the same in the ultrasound treatment group over the three time periods.

A simple bar graph like the one in figure 4a is easy to understand. It compares grip strength scores 1 to 3 in both treatment groups. This figure suggests that the mean grip strength increases, over time, in the OMT group while it decreases in the ultrasound group. To further explore this relationship a similar chart can be created, as in figure 4b, that incorporates the mean grip strength at time 2. Further subdivision of the mean grip strength values was done to examine the influence of both treatment group and gender; an example of this can be seen in figure 4c.

A scatterplot examining the interaction of treatment group and gender on grip strength is shown in figure 5a. This figure shows a seemingly equal distribution of strength values between both treatment groups. The scatterplot that shows the individual average grip strength values based on treatment group and gender (figure 5b) clearly shows that males tend to have higher grip strength values, no matter the treatment group, in comparison to females. Similar results were illustrated in the bar graphs. This observation may have implications for future conformational and inferential statistics; if average grip strength were to be assessed based only on treatment group there may not be

a significant difference between the two groups. If gender is also taken into consideration it is possible that significant differences between the groups may be found.

Two other scatterplots comparing the average grip strength with years since diagnosis (years affected with CTS) and functional status were also computed. The plot (figure 6) looking at average grip strength and years affected seems to show a slight negative log curve. Those who have most recently been diagnosed with CTS generally have higher grip strength scores in comparison to those who have suffered from the condition for a longer period of time. The plot (figure 7) using average grip strength and functional status scores indicates a slight possible negative sloping line though this pattern is not very distinct and may change as more patients are enrolled in the study.

B) Symptom Severity Scores

Figures 8 through 14 are visual representations of symptom severity scores. Some of these figures look at average symptom severity while others examine symptom severity scores at each time point. Symptom severity scores 1, 2 and 3 are representative of scores recorded at visits 3, 6 and 9 respectively, just as was seen in the grip strength scores.

Figures 8 and 8a are box plots of symptom severity scores 1,2 and 3. Figure 8a shows the overall symptom severity score for all subjects, while figure 8b separates the scores based on treatment group. Figure 8 shows that the symptom severity scores decrease over time. When these scores are separated based on treatment groups it is evident that symptom severity scores decrease in both groups over time. Though the OMT group's scores decrease continuously between each measurement time, while the

ultrasound group's scores decrease from time 1 to time 2 and then increase from times to 2 to 3, with a net decrease in symptom severity score.

Figures 9a and 9b are stem-and-leaf displays of the symptom severity scores. Figure 9a represents the scores separated into measurement times 1, 2 and 3, while figure 9b is the average symptom severity score. Figure 9a shows that symptom severity scores decrease from time 1 to time 3 and show the frequency of each score. Figures 10a and 10b are histograms that show the exact same data from the stem-and-leaf displays in the form of histograms. These graphs still show frequency of scores at the intervals but have the added advantage of more easily view the distribution of the data. The same observations that were drawn from the stem-and-leaf displays can be seen in the histograms. The histograms also allow one to note that each series of data is approximately normally distributed within the reported scores but are all negatively skewed in comparison to the overall range of possible scores.

Bar graphs of these distributions and relationships among symptom severity scores are shown in figures 11a, 11b and 11c. Figure 11a again shows that symptom severity scores decrease from time 1 to time 3 in both the OMT and ultrasound groups. Though the OMT groups seems to experience a more dramatic decrease in scores. The next figure (figure 11b) incorporates time measure 2 into the graph and shows that scores in the OMT group consistently decrease over time while the scores in the ultrasound group first decrease, then increase slightly but still result in an overall decrease in score over time. Figure 11c further subdivides the groups into treatment groups based on gender. This graph shows that OMT males and females as well as ultrasound females

experience a gradual decrease in symptom severity score over time. The ultrasound males actually experience the initial decrease and then a subsequent increase in scores between times 1 and 2, and 2 and 3, but then still show an overall decrease in score from time 1 to 3.

Scatterplots of symptom severity scores by treatment group, treatment group and gender, years since diagnosis and functional status are shown in figures 12a, 12b, 13 and 14. The first two scatterplots showing average symptom severity based on treatment group and treatment groups subdivided by gender seem to show no real patterns. Figure 13 shows symptom severity scores 1, 2 and 3 each versus years since diagnosis suggests that those with the longest time since diagnosis have slightly higher symptom severity scores compared to those with fewer years since diagnosis. Though this observation is not conclusive because the scores for those with the most years since diagnosis are no higher than some of the score experienced by individuals with a smaller number of years since diagnosis.

Figure 14 shows an interesting trend in comparing functional status scores at times 1, 2 and 3 with symptom severity scores 1, 2 and 3, respectively. The plots show a slightly positively sloping line, indicating a positive relation in that, as functional status scores increase, symptom severity scores also increase and vice versa (when functional status scores decrease, symptom severity scores also decrease).

C) Visual Analog Scores

The remaining figures (figures 15 through 21) deal with visual analog scale (VAS) scores, for level of self-reported pain, that were recorded at all six visits. Figures

15a, 15b and 15c are box plots of this outcome measure. Average pre- and post-VAS scores are shown based on treatment group in figure 15a. This plot shows that average VAS scores decreased from pre- to post-treatment in both groups, and slightly more so in the OMT group. One question that may be generated from this plot is what is the decrease in pre-test and post-test scores over treatment intervals. Figures 15b and 15c show pre-treatment scores (15b) and post-treatment scores (15c) for both groups. These plots both show that pre-treatment and post-treatment VAS scores decrease over treatment times in both the OMT and ultrasound groups.

Figure 16a shows the stem-and-leaf displays for average pre and post treatment VAS scores where figure 16b separates the stem-and-leaf displays based on pre and post treatment VAS scores at all six visits. The histograms in figures 17a and 17b present the same data as in the stem-and-leaf displays of figure 16a and 16b respectively.

A bar graph of average pre and post treatment VAS scores by treatment group is shown in figure 18a and it shows that there is a decrease in VAS score between before and after treatment in both the OMT and ultrasound group though this finding is seemingly greater in the OMT group. Figure 18b shows two bar graphs, the first shows the pre treatment VAS scores for all visits and the second shows post treatment VAS scores for each visit. The first graph clearly shows that VAS scores drop from visit 1 to visit 2 then gradually increase to visit 4 after which the scores drastically drop at visit 5 and slightly increase again at visit 6 with an overall decreasing trend in VAS scores. The second graph shows that post treatment VAS scores continuously exhibit a gradual decrease in score value from visit 1 to visit 3; at visit 4 the score increases slightly and

then gradually decreases to visit 6. These graphs were replicated in figure 18a for each treatment group, with similar patterns being expressed. The bar graph in figure 18c shows that all treatment groups separated by gender experience a decrease in the average VAS score from pre to post treatment. Ultrasound males have considerably higher pre- and post treatment VAS scores in comparison to the other three groups.

Scatterplots are shown in figures 19 through 21. The scatterplots in figure 19a show the average pre and post treatment VAS scores based on treatment group and those in figure 19b show the average pre and post treatment VAS scores based on treatment group by gender. There are no obvious patterns to these scatterplots. Figure 20 shows that there may be a slight negative log curve relationship (in which as years since diagnosis increase the VAS score decreases) between years since diagnosis and pre treatment VAS scores, but not the post-treatment VAS scores. The two scatterplots in figure 21 both suggest a positively sloped linear relationship between functional status and both (average) pre and post treatment VAS scores.

OB

Overview of EDA Applications

The OB study collected a very large number of outcome measures. A particularly large number of EDA analyses could be performed with this data. Patient history information about tobacco, alcohol and drug use could be used to see if there may be a relationship to low back pain during pregnancy, and whether it is related to any of the measures recorded during delivery. Levels of low back pain measured with the Roland-Morris could be plotted with patients' confidence in treatment scores. Other analyses

could be done that directly relate to the hypotheses of this study. In this way, the researchers would gain insights into the data prior to conducting significance testing.

DISCUSSION

OMM physician researchers face particular challenges in conducting and analyzing their research data due to several factors. Because of the physiological or anatomical changes that manual medicine and OMT produce, and the limited availability of validated and reliable objective measures to assess the outcomes of specific interventions, researchers are left to evaluate much of their interventions on patient-centered subjective outcomes. There is a general bias that exists among the scientific community that places a higher degree of confidence in objective, electronically or physically obtained outcomes versus those reported by the patient. Some believe this bias toward solidly objective measures to be unjustified, and they suggest that patient generated outcomes such as quality of life and level of pain are just as, if not more reliable, than those objective measures (Kane, 1997).

The best OMM studies use a combination of subjective and objective measures. Investigators continue to explore new methods of measuring outcomes if existing measures do not adequately answer the research question. Unfortunately existing studies of the efficacy of OMM for the same disease processes have not used the same outcome measures. Also, few researchers have been faithful to one clinical focus so that information from their studies has increased over time. If one article or abstract is published showing "no effects," no one seems to study it further. This is true despite the

small numbers of subjects in most trials, and some inadequacies in the research designs. Limited collaboration with basic scientists, limited to no description of the data collected, and the absence of large funded randomized and controlled clinical trials. Replications of studies have not been done to strengthen the design or power. Placebo-potency or placebo-effects have not been adequately explored. Analysis of within group and between group outcomes has not been adequately explored. As a consequence, there seems to be no preferred set of outcome measures.

Each published OMM study is different. Primary outcomes seem to be derived from the researchers' hunches rather than previously published research, making it difficult to learn and devise new studies based on previous research. It is unclear as to the rigor of the definition of the clinical population being studied (inclusion and exclusion criteria) or the full measure of "control" used in control groups that might strengthen the reliability of any findings associated with the treatment under study.

The problems concerned with generating consistency of design and continuing a focus and theme in OMM research are two of the reasons the ORC was created. Another contributing factor to the problems with the research and outcomes being selected in OMM research may be related to limited research training in the education of Osteopathic medical students. Although there are a few fellowships in OMM, most focus on teaching and clinical skills, and provide no research training or experience. Hence some osteopathic physicians receive no education as to the importance of research. This can be limiting to the true potential of the profession in a number of ways.

1) Those not educated in the importance of research may not believe it is important to validate the practices of OMT; and

2) They may not carry on the practice of performing OMT in the clinical setting because they do not understand its potential value in the treatment of patients'.

This under values the philosophical principles of the profession. OMM is one of the most distinctive features of osteopathic medicine. If osteopathic physicians do not understand and utilize these techniques than one of the profession's greatest assets could be lost. Patients would miss out on a possible treatment that is effective.

If a researcher in an exploratory study finds results that are shown to be statistically insignificant with a p-value greater than 0.05 he or she should ask what actually occurred in the data. This could be a consequence of one of the following:

a) A poor research design;

b) Less-than-optimal statistical techniques;

c) Small sample size;

d) Violating the assumptions of a particular statistical test; and

e) Heavy reliance on subjective outcome measures.

SUMMARY AND RECOMMENDATIONS

This paper began by explaining the relevancy of this topic and my interest in pursuing this topic of study. Then I discussed the history and relevant aspects of the osteopathic profession and OMM and the current published research. Types of outcome measurements were discussed, as were the various types of clinical outcomes used in

clinical trials of OMM. A history of EDA was given, as was a discussion of its relevancy to OMM research. Some of the most frequently used tools of EDA found in the literature were discussed and subsequently applied to data from two active studies.

This project had four steps in achieving the aims:

1) Identify and describe the most frequently used outcome measures described in OMM research published since 1993;
2) Evaluate outcomes used in the published research of the profession;
3) Describe EDA and its relevance to the types of data used in OMM research; and
4) Illustrate the ways EDA can be used in two current studies to gain insights into the data and shape future research questions.

This paper is intended to provide recommendations for future OMM researchers.

The most important aspect of EDA is not the techniques used but the mindset for approaching data analysis. Some authors suggest that EDA should be used strictly on preliminary data that is being used to develop a hypothesis or theory and that a completely new set of data must be used to prove these hypotheses or theories (Behrens, 1997). Other authors believe that EDA can be used as a tool in conjunction with confirmatory and inferential statistics to first understand the nature of the data and generate questions to be answered by the subsequent statistical analyses (Lederman, 1992; Ferketich & Verrans, 1986). I suggest that EDA can be used in either of these two capacities.

By using the EDA in conjunction with the final statistical analysis it may be possible for the researcher to obtain far better insights into the results of their clinical

trials. They can answer the question why this result occurred in more ways than with hypothesis testing alone.

Research on the mechanisms and efficacy of OMT have a history of being anecdotal and focused mainly on case studies, or underpowered with inadequate rigor in the research design. Insurance companies tend not to cover treatments that only make a person feel and function better. Third party payers require scientific evidence. Hence comes the term "evidence-based-medicine." Nonetheless, we all need to learn to crawl before we walk, and OMM researchers are learning how to incorporate more objective physiologic and anatomic measures in their research designs along with subjective outcome measures.

The strategies of EDA may be of particular usefulness to clinical research investigators because they allow the researcher to examine the data prior to any hypothesis testing. Using EDA to examine data may surface or reveal possible relationships among different variables that might better guide the statistical tests; it also might help elucidate the data behind "conclusions." These procedures may be particularly useful when examining outcomes data from clinical science versus those from the basic sciences because clinical outcomes tend not to be interval or ratio. The Osteopathic Research Center is working to foster research initiatives that bridge the gap between the basic and clinical sciences, and subsequently improve the ways researchers use and analyze objective and subjective outcome measures.

Medical school and postgraduate research fellows conducting preliminary clinical trials at the ORC have only basic statistical training in one statistics course. They design

their studies with mentors and tend to rely heavily on traditional, hypothesis-testing statistical procedures. Most of these students conduct small trials with small numbers of subjects and have broad, exploratory aims to their research. The aim of many of these projects is to gain a better understanding into the factors that may potentially be related to the efficacy of OMT as a treatment for a specific disease or condition. From my experience working with these students, some of their challenges lie in:

1) Inability because of the state of the science to be precise and narrow in formulating the research questions;
2) Need to have more science in selecting outcome measures to answer the research questions;
3) Need more attention to the design of the research in relationship to the data to be collected;
4) Need more consultation on statistical techniques that could be used to analyze the data;
5) Need more support to interpret the results obtained from the data analysis procedures; and
6) Need to understand how to address statistically insignificant but possibly clinically important findings.

EDA is uniquely applicable to the generation of hypotheses and theories because it provokes the investigator to ask more questions while providing a tool to understand each variable as well as the interactions among variables in thorough detail. I also find EDA to be a valuable descriptive tool that can be used in preliminary and pilot studies

with a simple design intended to try to answer very broad questions and possibly generate hypotheses for future research. The use of EDA may be very important in the understanding of primary as well as secondary research questions in these types of studies. If one tries to examine the dichotomous roles of EDA described in the literature and apply the idea expressed in this paper, it seems that both described uses of EDA are applicable. EDA can be used to both to generate hypotheses and can be used in conjunction with conformational statistics to adequately answer postulated questions.

APPENDIX

Table 1: Outcome Measures used in OMT research

Study Authors	Investigated OMT and ...	Outcome measures utilized
Licciardone	Low back pain	• demographic information • SF-36 • VAS • Roland-Morris • number of lost work/school days • patient satisfaction
Andersson	Low back pain	• demographic information • VAS • Roland-Morris • Oswestry • N. Amer. Spine Society Outcome Assess. Instrument • patient's acceptance of pain • patient pain drawing • range of motion • degree of straight leg raise
Noll	Elderly hospitalized pneumonia patients	• demographic information • Simplified Acute Physiology Score • vital signs (temperature, pulse, respiratory rate) • X-rays
Gamber	Fibromyalgia	• tender points • Chronic Pain Experience Inventory • Present Pain Intensity Rating Scale • self-evaluation questionnaire • Stanford Arthritis Center Disability and Discomfort Scales: Health Assessment Questionnaire • Center for Epidemiological Studies Depression Scale
Jarski	Recovery from surgery	• days to independent negotiation of stairs • distance ambulated • intramuscular analgesic use • length of hospital stay • patient's perception of treatment
Noll	Elderly pneumonia patients	• IV use • antibiotic use • length of hospital stay
Handler	Headache pain relief	• assessment of pain and symptoms • medication use
Smutny	Parkinson's Disease	• 3-D gait analysis • general fitness • cardiovascular fitness • general health • osteopathic structural analysis • nutrition assessment • assessment of home environment
Brittain and Scandalis	Post-hip and knee surgery recovery	• standard disability measurement • Functional Independence Measure
Steele	Otitis media	• demographic information • eight question behavior assessment questionnaire • tympanometric measures (adequacy, baseline adequacy peak, VEA, TPP and classification) • number of office visits for otitis media • number of prescriptions dispensed • surgical referral

Table 2: Outcome Measures Used in CTS and OB Studies

Outcomes used in the CTS study:

- Demographic information
- Strength measures
  - Grip
  - Key pinch
  - Tripod pinch
  - Tip pinch
- Symptom severity
- Functional status
- Nerve conduction studies:
  - Median, motor and sensory, latency and amplitude
  - Ulnar, motor and sensory, latency and amplitude
  - Difference in median/ulnar motor latency
  - Difference in median/ulnar sensory latency
- MRI data:
  - Tunnel AP diameter
  - Transverse diameter
  - Cross-sectional area
  - Amount of marrow
  - Carpal tunnel mean signal
  - Nerve rect. Signal
  - Mean nerve geog.

Outcomes used in the OB study:

- Demographic information
- Confidence in Treatment questionnaire for both OMT and standard care
- Roland-Morris
- Quadruple VAS
- Sv12v2
- Hospital admission information
- Delivery information

EDA FIGURES

Figure 1a: Box plot of grip strengths 1, 2 and 3

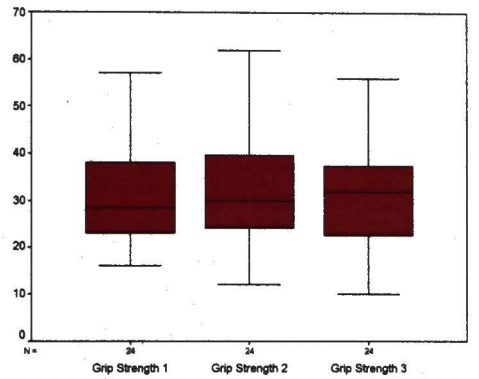


Figure 1b: Box plot of grips strengths 1, 2 and 3 based on treatment group

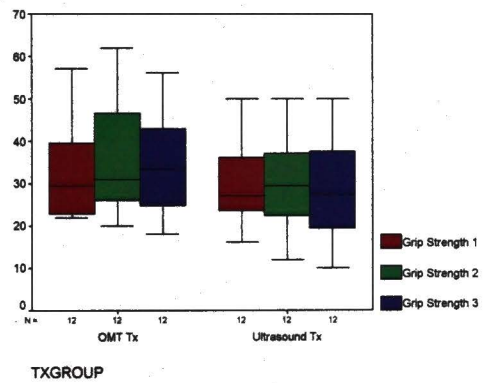


Figure 2a: Stem-and-leaf display of average grip strength

average grip strength

```
average grip strength Stem-and-Leaf Plot

 Frequency     Stem &  Leaf

     1.00        1 .  2
     1.00        1 .  6
     5.00        2 .  00233
     4.00        2 .  6669
     6.00        3 .  012234
     2.00        3 .  78
      .00        4 .
     2.00        4 .  89
     2.00        5 .  02
     1.00        5 .  8

 Stem width:     10.00
 Each leaf:       1 case(s)
```

Figure 2b: Stem-and-leaf displays of grip strengths 1, 2 and 3

Grip Strength 1

Grip Strength 1 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
2.00	1 .	68
5.00	2 .	22224
6.00	2 .	566889
4.00	3 .	0022
1.00	3 .	6
2.00	4 .	03
.00	4 .	
3.00	5 .	000
1.00	5 .	7

Stem width: 10.0
Each leaf: 1 case(s)

Grip Strength 2

Grip Strength 2 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
2.00	1 .	24
9.00	2 .	013448899
7.00	3 .	0022468
1.00	4 .	1
4.00	5 .	0024
1.00	6 .	2

Stem width: 10.0
Each leaf: 1 case(s)

Grip Strength 3

Grip Strength 3 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
1.00	1 .	0
3.00	1 .	889
4.00	2 .	0144
2.00	2 .	56
5.00	3 .	01344
4.00	3 .	5678
.00	4 .	
1.00	4 .	9
3.00	5 .	004
1.00	5 .	6

Stem width: 10.0
Each leaf: 1 case(s)

Figure 3a: Histogram of average grip strength

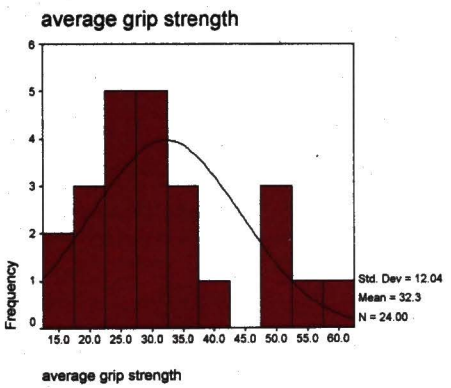


Figure 3b: Histograms of grip strengths 1, 2 and 3 based on treatment group

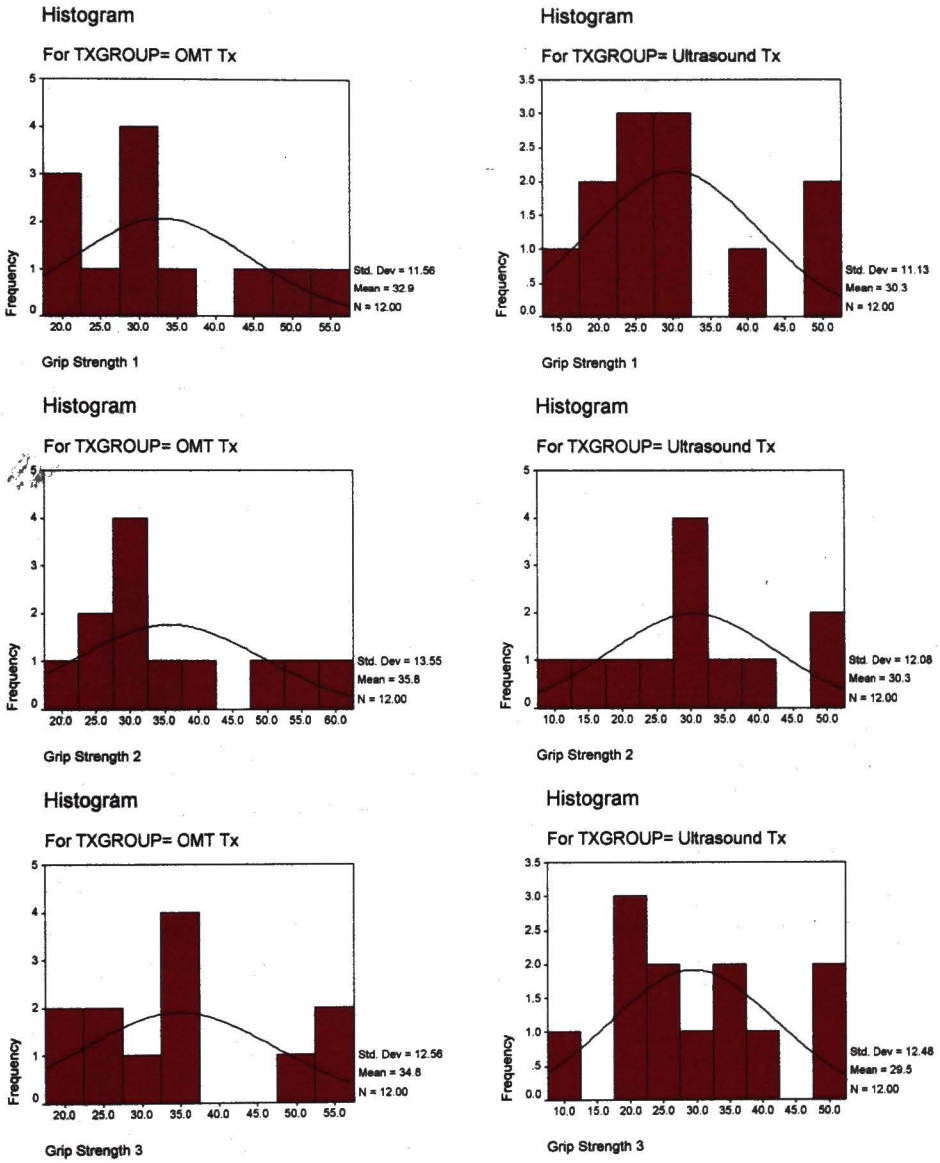


Figure 4a: Bar graph of grip strengths 1 and 3 based on treatment group

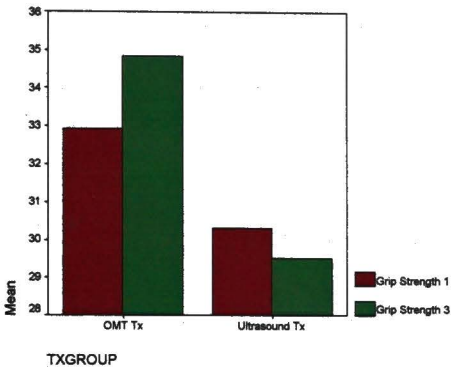


Figure 4b: Bar graph of grip strengths 1, 2 and 3 based on treatment group

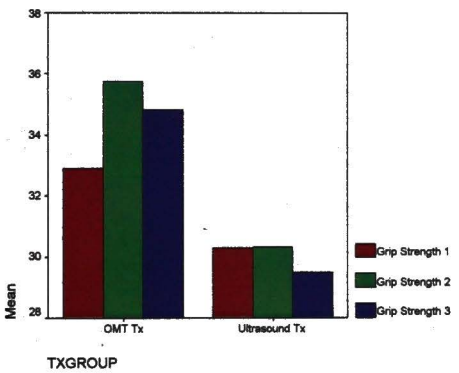


Figure 4c: Bar graph of grip strengths 1, 2 and 3 based on treatment group and gender

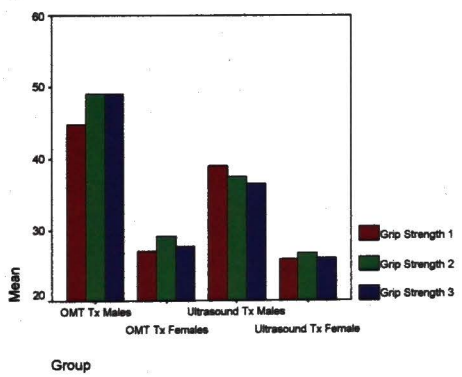


Figure 5a: Scatterplot of average grip strength based on treatment group

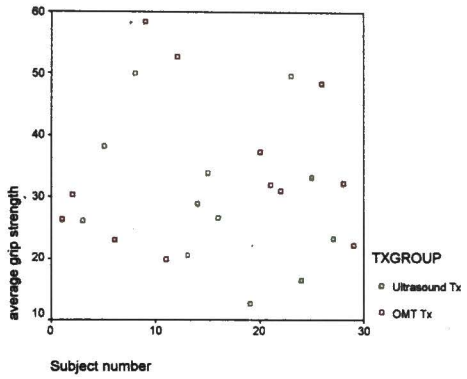


Figure 5b: Scatterplot of average grip strength based on treatment group and gender

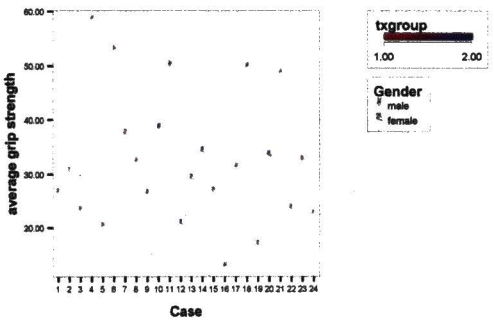


Figure 6: Scatterplot of average grip strength and years since diagnosis

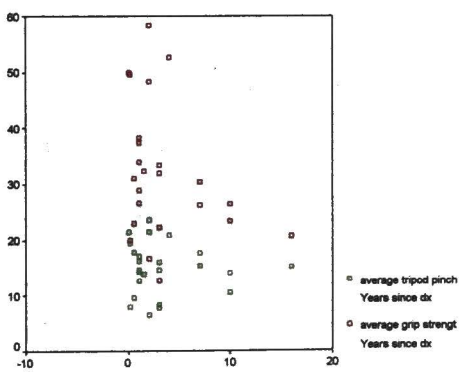


Figure 7: Scatterplot of average grip strength and functional status

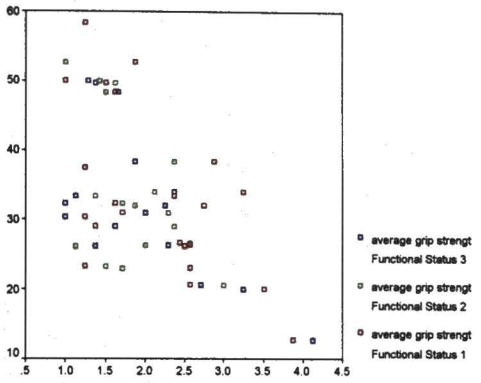


Figure 8a: Box plot of symptom severity scores 1, 2 and 3

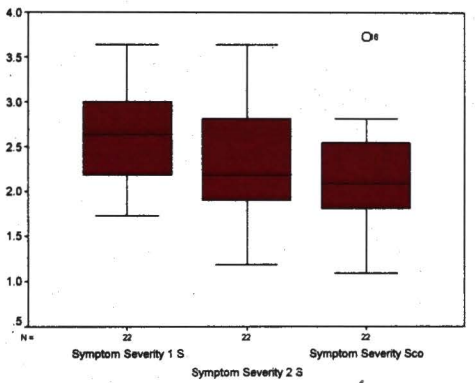


Figure 8b: Box plot of symptom severity scores 1, 2 and 3 based on treatment group

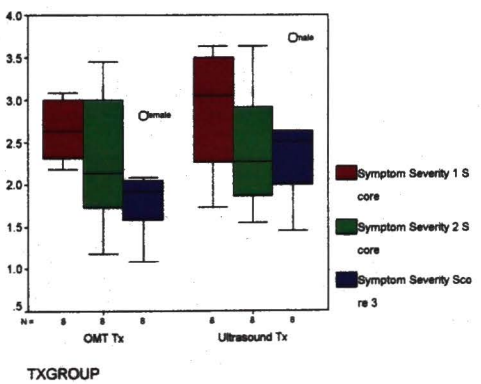


Figure 9a: Stem-and-leaf displays of symptom severity scores 1, 2 and 3

Symptom Severity 1 Score

Symptom Severity 1 Score
Stem-and-Leaf Plot

Frequency	Stem &	Leaf
1.00	1 .	7
7.00	2 .	
0001144		
4.00	2 .	5789
4.00	3 .	0004
2.00	3 .	56

Stem width: 1.00
Each leaf: 1 case(s)

Symptom Severity 2 Score

Symptom Severity 2 Score
Stem-and-Leaf Plot

Frequency	Stem &	Leaf
1.00	1 .	1
5.00	1 .	55999
6.00	2 .	
001113		
2.00	2 .	88
3.00	3 .	124
1.00	3 .	6

Stem width: 1.00
Each leaf: 1 case(s)

Symptom Severity Score 3

Symptom Severity Score 3
Stem-and-Leaf Plot

Frequency	Stem &	Leaf
1.00	1 .	0
.00	1 .	
2.00	1 .	44
1.00	1 .	7
1.00	1 .	8
6.00	2 .	
000001		
1.00	2 .	2
2.00	2 .	45
2.00	2 .	66
1.00	2 .	8
1.00 Extremes (>=3.7)		

Stem width: 1.00
Each leaf: 1 case(s)

Figure 9b: Stem-and-leaf display of average symptom severity score

average symptom severity score

average symptom severity score Stem-and-Leaf Plot

Frequency	Stem &	Leaf
5.00	1 .	66999
9.00	2 .	000122233
5.00	2 .	66688
2.00	3 .	01
1.00	3 .	6

Stem width: 1.00
Each leaf: 1 case(s)

Figure 10a: Histogram of average symptom severity score

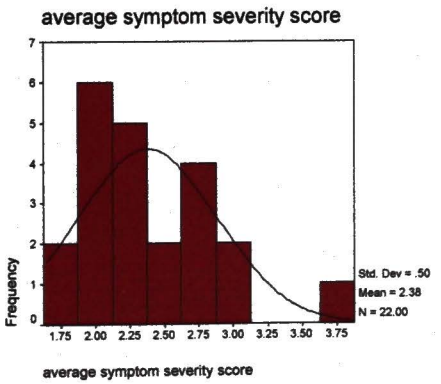
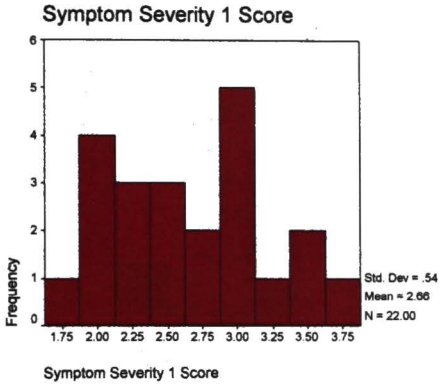
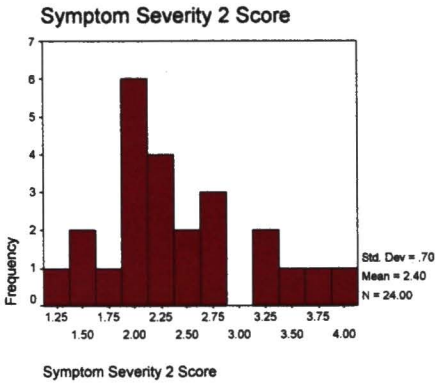


Figure 10b: Histograms of symptom severity scores 1, 2 and 3





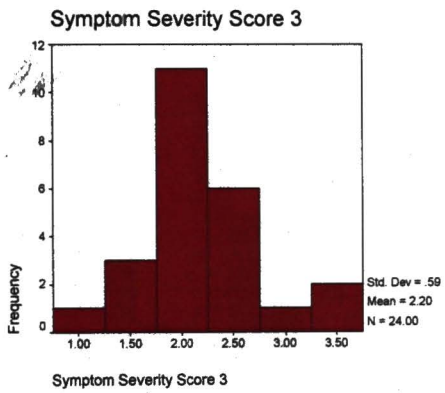


Figure 11a: Bar graph of symptom severity scores 1 and 3 based on treatment group

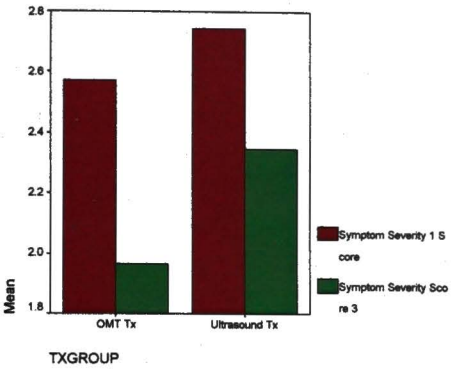


Figure 11b: Bar graph of symptom severity scores 1, 2 and 3 based on treatment group

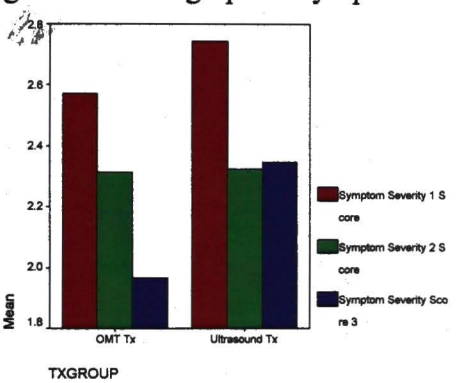


Figure 11c: Bar graph of symptom severity scores 1, 2 and 3 based on treatment group and gender

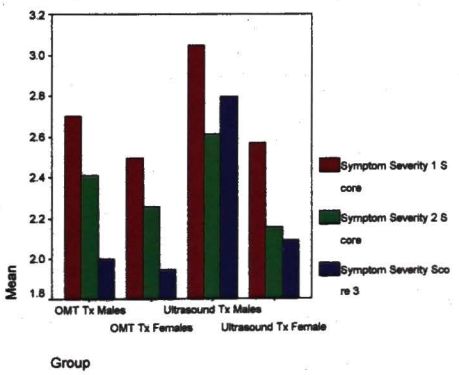


Figure 12a: Scatterplot of average symptom severity score based on treatment group

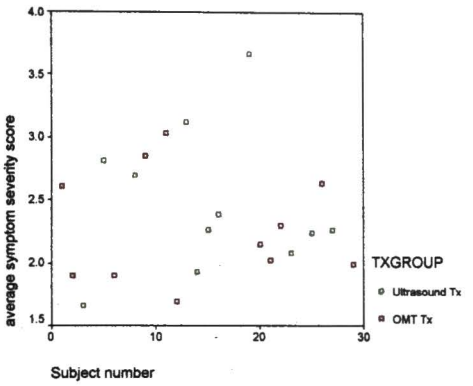


Figure 12b: Scatterplot of average symptom severity score based on treatment group and gender

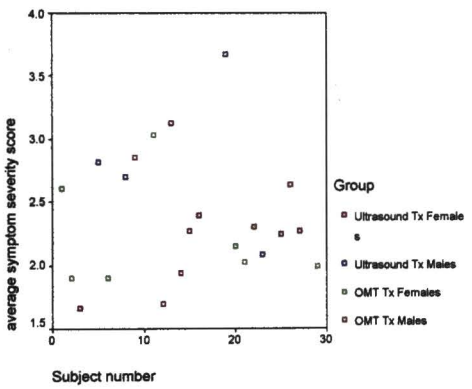


Figure 13: Scatterplot of symptom severity scores 1, 2 and 3 based on years since diagnosis

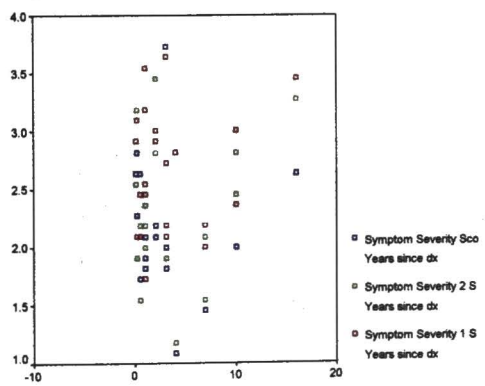


Figure 14: Scatterplot of symptom severity scores 1, 2 and 3 based on function status scores 1, 2 and 3

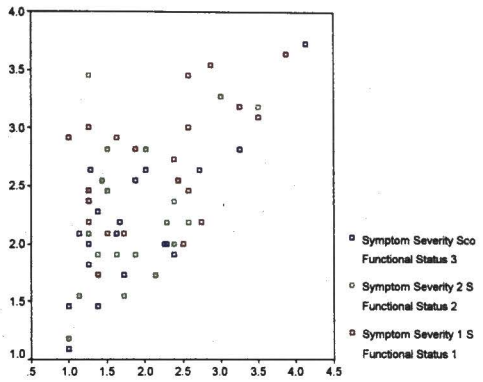


Figure 15a: Box plot of average pre- and post-VAS based on treatment group

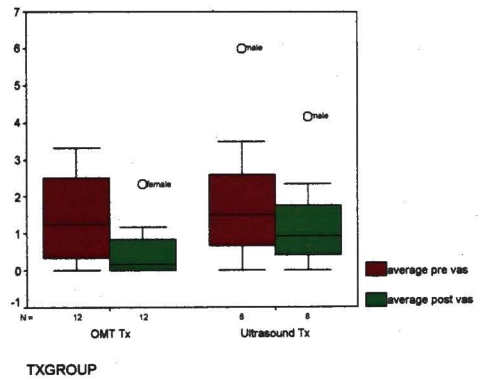


Figure 15b: Box plot of pre-VAS scores 1–6 based on treatment group

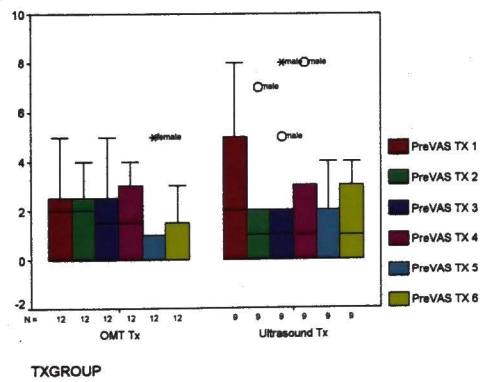


Figure 15c: Box plot of post-VAS scores 1-6 based on treatment group

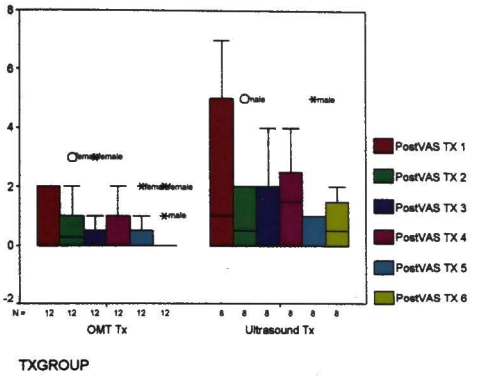


Figure 16a: Stem-and-leaf displays of average pre- and post-VAS scores

average pre vas

average pre vas Stem-and-Leaf Plot

Frequency	Stem &	Leaf
8.00	0 .	00015666
5.00	1 .	05556
.00	2 .	
4.00	3 .	0035
.00	4 .	
.00	5 .	
1.00	6 .	0

Stem width: 1.00
Each leaf: 1 case(s)

average post vas

average post vas Stem-and-Leaf Plot

Frequency	Stem &	Leaf
9.00	0 .	000000123
3.00	0 .	666
4.00	1 .	0111
.00	1 .	
1.00	2 .	3
1.00 Extremes		(>=4.2)

Stem width: 1.00
Each leaf: 1 case(s)

Figure 16b: Stem-and-leaf displays of pre-VAS scores 1-6

PreVAS TX 1

PreVAS TX 1 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
8.00	0 .	00000000
1.00	1 .	0
3.00	2 .	000
1.00	3 .	0
1.00	4 .	0
2.00	5 .	00
1.00	6 .	0
.00	7 .	
1.00	8 .	0

Stem width: 1.0
Each leaf: 1 case(s)

PreVAS TX 2

PreVAS TX 2 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
7.00	0 .	0000000
.00	0 .	
2.00	1 .	00
.00	1 .	
6.00	2 .	000000
.00	2 .	
2.00	3 .	00
1.00 Extremes		(>=7.0)

Stem width: 1.0
Each leaf: 1 case(s)

PreVAS TX 3

PreVAS TX 3 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
9.00	0 .	000000000
2.00	1 .	00
2.00	2 .	00
1.00	3 .	0
.00	4 .	
3.00	5 .	000
1.00 Extremes		(>=8.0)

Stem width: 1.0
Each leaf: 1 case(s)

PreVAS TX 4

PreVAS TX 4 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
7.00	0 .	0000000
.00	0 .	
2.00	1 .	00
.00	1 .	
1.00	2 .	0
.00	2 .	
5.00	3 .	00000
.00	3 .	
2.00	4 .	00
1.00 Extremes		(>=8.0)

Stem width: 1.0
Each leaf: 1 case(s)

PreVAS TX 5

PreVAS TX 5 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
10.00	0 .	0000000000
.00	0 .	
4.00	1 .	0000
.00	1 .	
1.00	2 .	0
3.00 Extremes		(>=4.0)

Stem width: 1.0
Each leaf: 1 case(s)

PreVAS TX 6

PreVAS TX 6 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
9.00	0 .	000000000
.00	0 .	
3.00	1 .	000
.00	1 .	
2.00	2 .	00
.00	2 .	
2.00	3 .	00
.00	3 .	
2.00	4 .	00

Stem width: 1.0
Each leaf: 1 case(s)

Figure 16c: Stem-and-leaf displays of post-VAS scores 1-6

PostVAS TX 1

PostVAS TX 1 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
10.00	0 .	0000000000
.00	0 .	
1.00	1 .	0
.00	1 .	
4.00	2 .	0000
3.00 Extremes		(>=5.0)

Stem width: 1.0
Each leaf: 1 case(s)

PostVAS TX 2

PostVAS TX 2 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
10.00	0 .	0000000000
1.00	0 .	5
3.00	1 .	000
.00	1 .	
3.00	2 .	000
1.00 Extremes		(>=5.0)

Stem width: 1.0
Each leaf: 1 case(s)

PostVAS TX 3

PostVAS TX 3 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
13.00	0 .	0000000000000
.00	0 .	
2.00	1 .	00
.00	1 .	
2.00	2 .	00
1.00 Extremes		(>=4.0)

Stem width: 1.0
Each leaf: 1 case(s)

PostVAS TX 4

PostVAS TX 4 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
10.00	0 .	0000000000
.00	0 .	
4.00	1 .	0000
.00	1 .	
2.00	2 .	00
2.00 Extremes		(>=3.0)

Stem width: 1.0
Each leaf: 1 case(s)

PostVAS TX 5

PostVAS TX 5 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
13.00	0 .	0000000000000
.00	0 .	
3.00	1 .	000
.00	1 .	
1.00	2 .	0
1.00 Extremes		(>=5.0)

Stem width: 1.0
Each leaf: 1 case(s)

PostVAS TX 6

PostVAS TX 6 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
13.00	0 .	0000000000000
.00	0 .	
3.00	1 .	000
.00	1 .	
2.00	2 .	00

Stem width: 1.0
Each leaf: 1 case(s)

Figure 17a: Histograms of average pre- and post-VAS scores

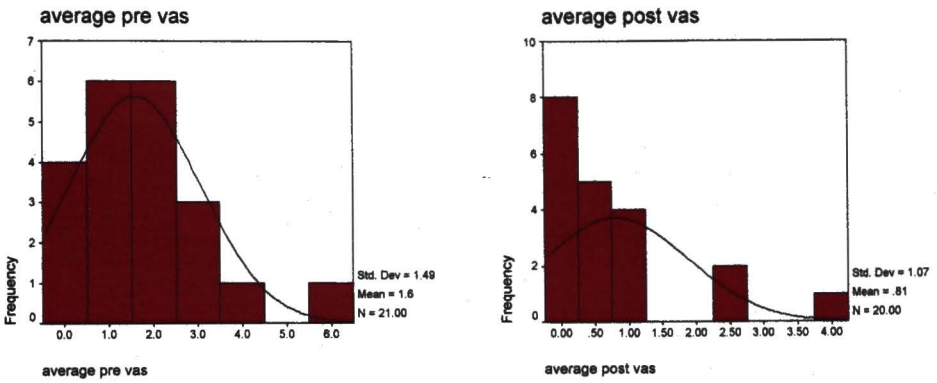
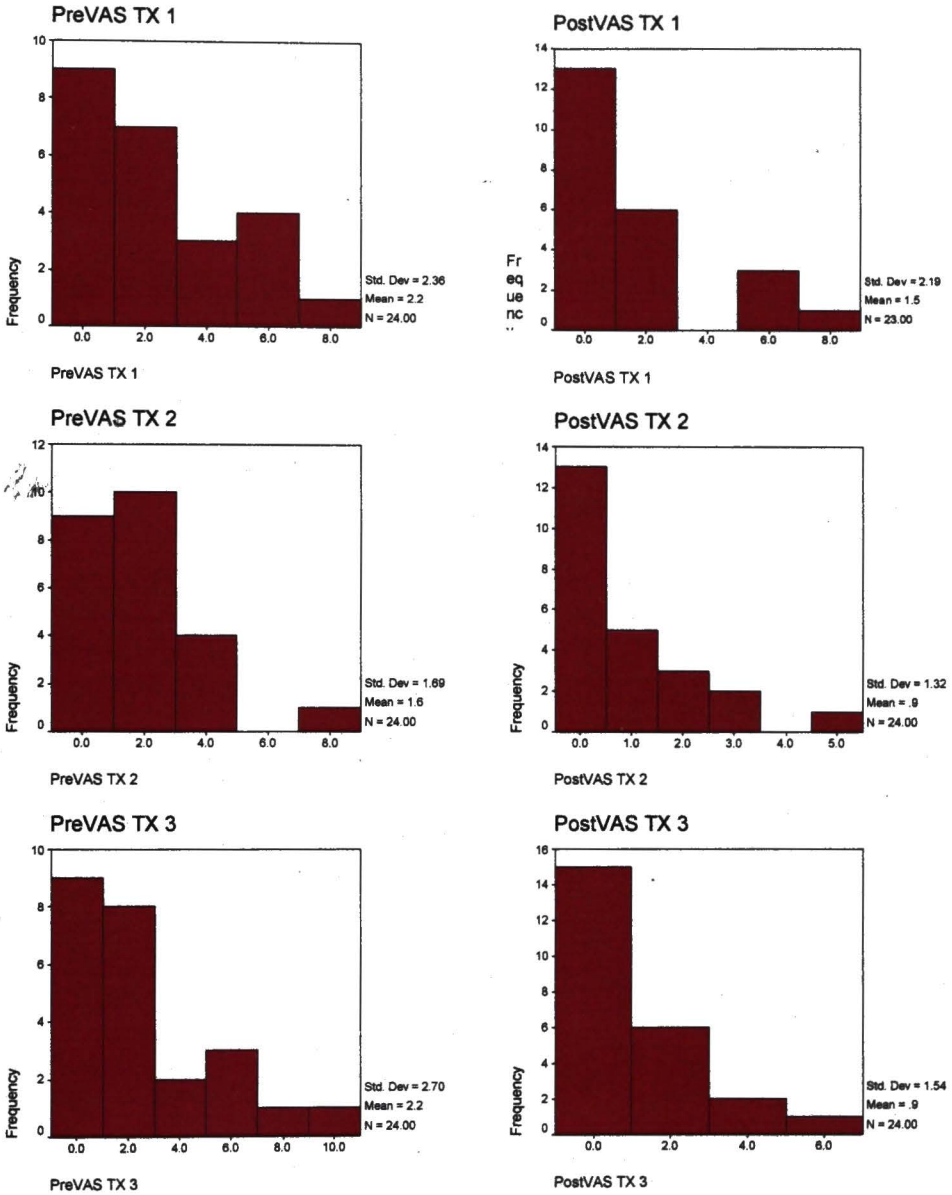
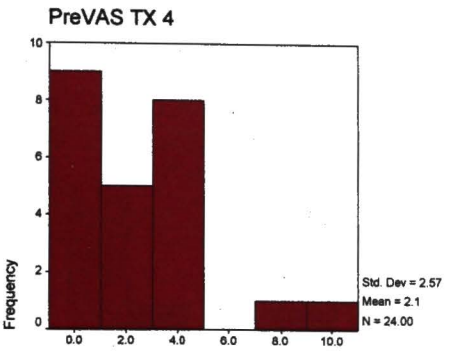
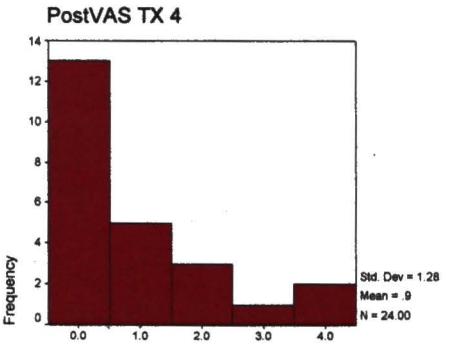
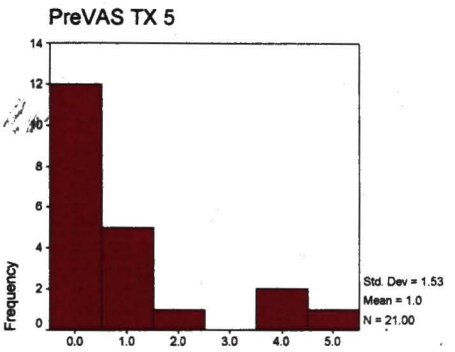


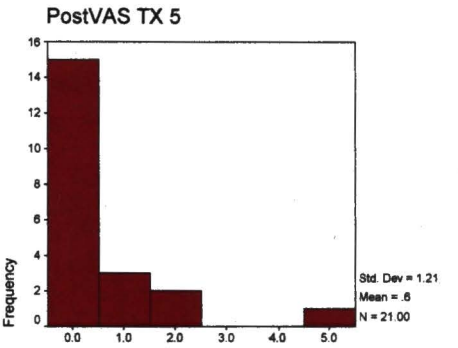
Figure 17b: Histograms of pre- and post-VAS scores 1-6

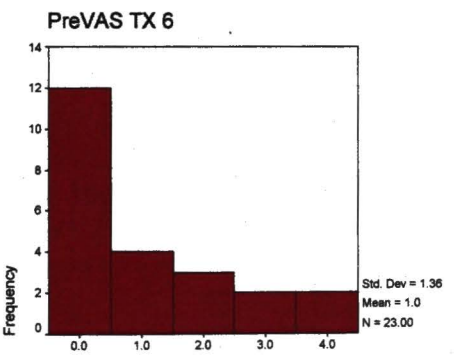



PreVAS TX 4
Frequency
Std. Dev = 2.57
Mean = 2.1
N = 24.00
PreVAS TX 4


PostVAS TX 4
Frequency
Std. Dev = 1.28
Mean = .9
N = 24.00
PostVAS TX 4


PreVAS TX 5
Frequency
Std. Dev = 1.53
Mean = 1.0
N = 21.00
PreVAS TX 5


PostVAS TX 5
Frequency
Std. Dev = 1.21
Mean = .6
N = 21.00
PostVAS TX 5


PreVAS TX 6
Frequency
Std. Dev = 1.36
Mean = 1.0
N = 23.00
PreVAS TX 6

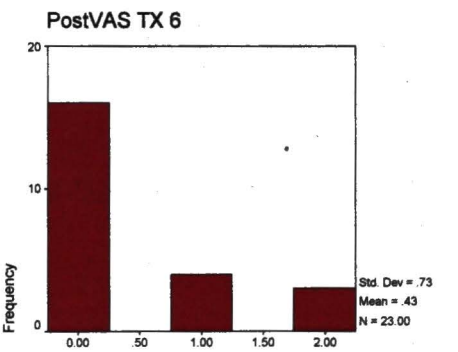

PostVAS TX 6
Frequency
Std. Dev = .73
Mean = .43
N = 23.00
PostVAS TX 6

Figure 18a: Bar graph of average pre- and post-VAS scores based on treatment group

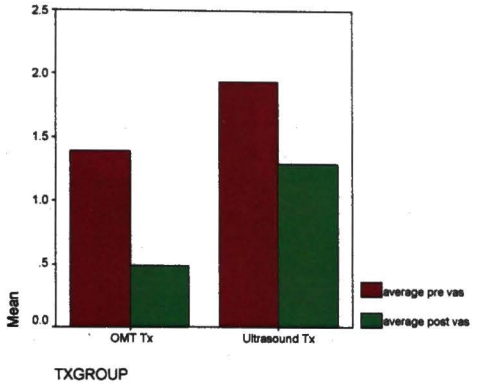


Figure 18b: Bar graphs of pre- and post-VAS scores 1-6

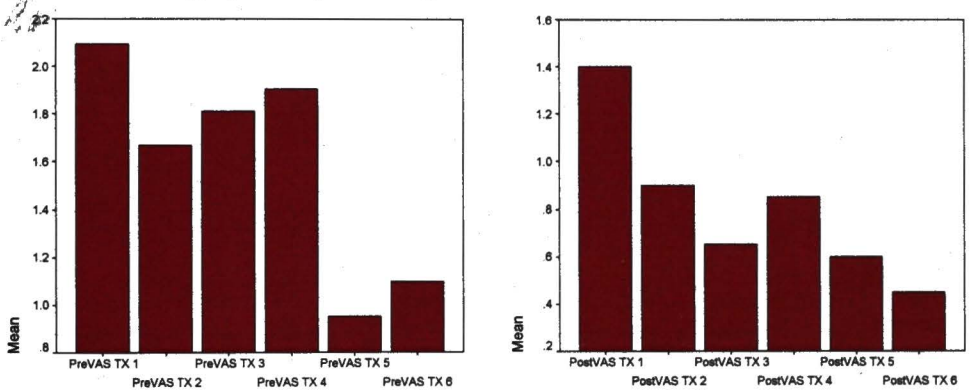


Figure 18c: Bar graph of average pre- and post-VAS scores based on treatment group and gender

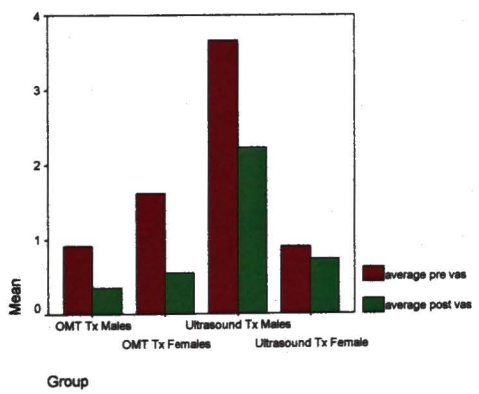


Figure 18d: Bar graphs of pre- and post-VAS scores 1-6 based on treatment group

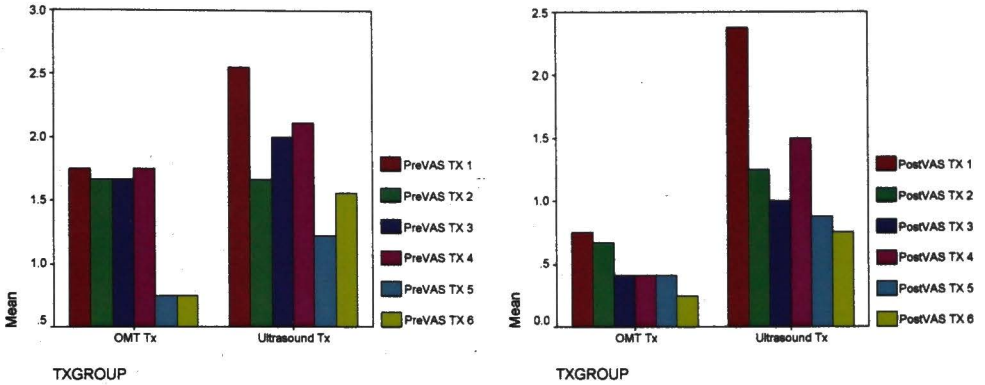


Figure 19a: Scatterplots of average pre- and post-VAS scores based on treatment group

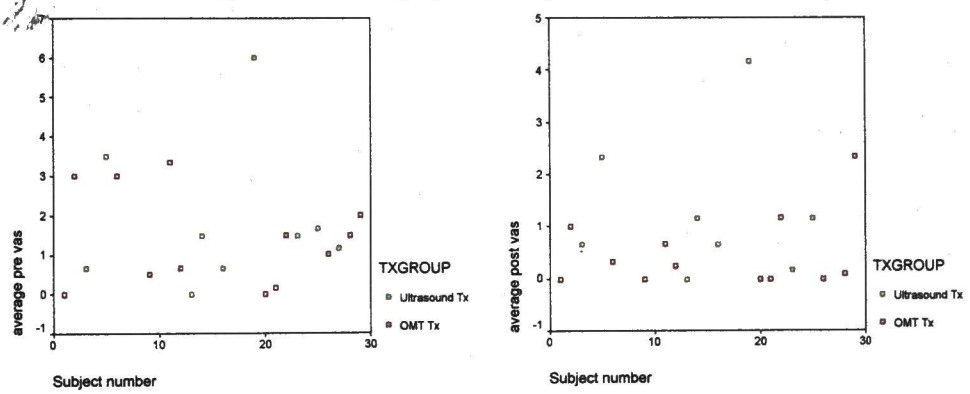


Figure 19b: Scatterplots of average pre- and post-VAS scores based on treatment group and gender

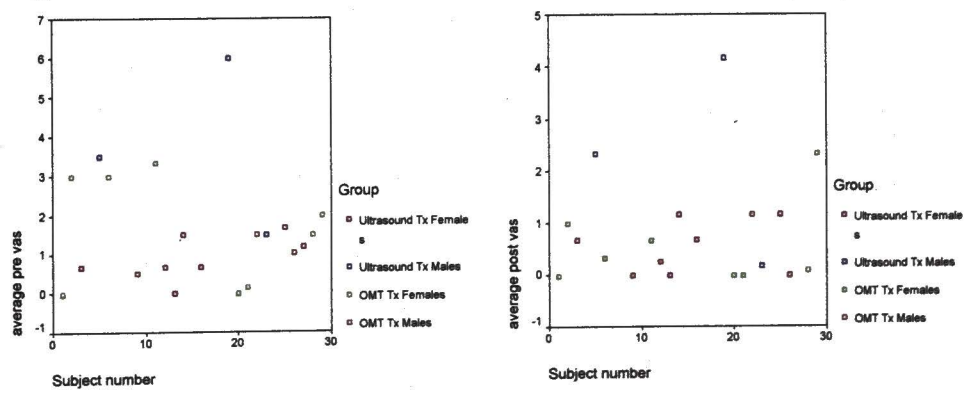


Figure 20: Scatterplot of average pre- and post-VAS scores and years since diagnosis

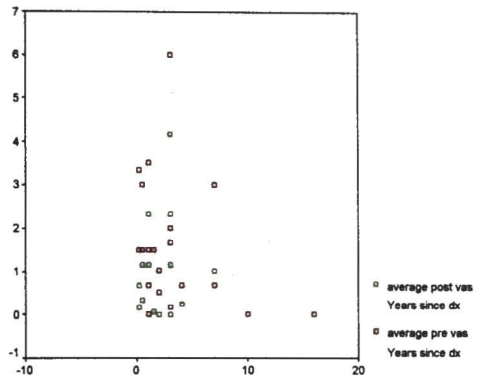
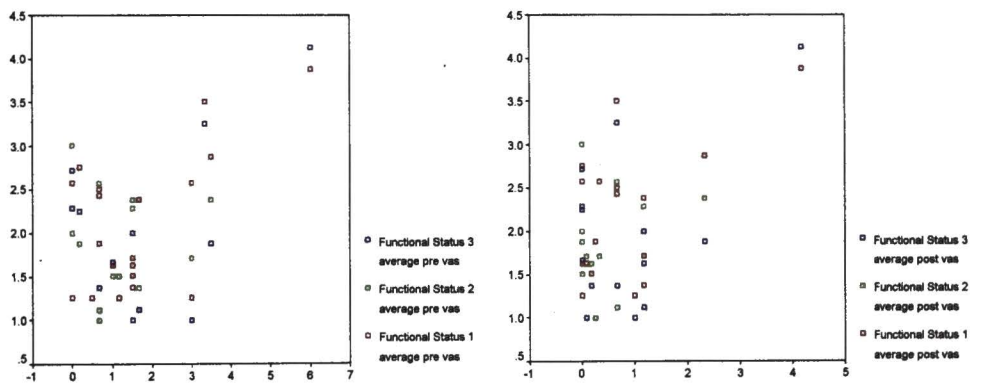


Figure 21: Scatterplots of average pre- and post-VAS scores and functional status scores 1, 2 and 3



REFERENCES

Abelson, R.P. (1995). *Statistics as principled argument.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Andersson, G.B.J., Lucente, T., Davis, A.M., Kappler, R.E., Lipton J.A. & Leurgans, S. (1999). A comparison of osteopathic spinal manipulation with standard care for patients with low back pain. *The New England Journal of Medicine, 341*(19), 1426-1431.

Behrens, J.T. (1997). Principles and procedures of exploratory data analysis. *Psychological Methods, 2*(2), 131-160.

Brittain, P.D., Stoll, S.T., Dickey, J.A., Romanski, V.V. & Licciardone, J.C. (1997). Efficacy of osteopathic manipulative treatment in improving clinical outcomes in patients with orthopedic diagnoses admitted to a hospital-based rehabilitation unit (Abstract). *The Journal of the American Osteopathic Association, 97*(8), 483.

Denis, D.J. (2003). *Theory & Science: Alternatives to null hypothesis significance testing.* Retrieved from the World Wide Web July 15, 2004. http//:theoryandscience.icaap.org/content/vol4.1/02_denis.html.

DiGiovanna, E.L. & Schiowitz, S. (1991). *An osteopathic approach to diagnosis and treatment.* Philadelphia: J.B. Lippincott Company.

Ferketich, S. & Verran, J. (1986). Exploratory data analysis: Introduction. *Western Journal of Nursing Research, 8*(4), 464-466.

Froehlich, G.W. (2004). *What is the chance that this study is clinically significant? A proposal for Q values.* Retrieved from the World Wide Web, July 15, 2004. http://www.acponline.org/journals/ecp/sepoct99/froehlich.htm. Originally published in Effective Clinical Practice, Sept/Oct. 1999.

Gamber, R.G., Shores, J.H., Russo, D.P., Jimenez, C., & Rubin, B.R. (2002). Osteopathic manipulative treatment in conjunction with medication relieves pain associated with fibromyalgia syndrome: Results of a randomized clinical pilot project. *The Journal of the American Osteopathic Association, 102*(6), 321-325.

Handler, P.J., Nicholas, A.S., Stewart, P.E., Goldstein, F.J. & Coughlin, P.M. (1998). Evaluation of the effect of osteopathic manipulative treatment on headache pain: Duration of pain relief (Abstract). *The Journal of the American Osteopathic Association, 98*(7), 390.

Hartwig, F. & Dearing, B.E. (1979). *Exploratory data analysis*. Newbury Park: Sage university Paper.

Hulley, S.B., Cummings, S.R., Browner, W.S., Grady, D., Hearst, N. & Newman, T.B. (2001). *Designing clinical research, 2nd ed.* Philadelphia: Lippincott Williams & Wilkins.

Jarski, R.W., Loniewski, E.G., Williams, J., Bahu, A., Shafinia, S., Gibbs, K. & Muller, M. (2000). The effectiveness of osteopathic manipulative treatment as complementary therapy following surgery: A prospective, match-controlled outcome study (Abstract). *Alternative Therapies in Health and Medicine, 6*(5), 77-81.

Johnson, S.M. & Kurtz, M.E. (2002). Conditions and diagnoses for which osteopathic primary care physicians and specialists use osteopathic manipulative treatment. *The Journal of the American Osteopathic Association, 102*(10), 527-540.

Jones, B.E. (1978). *The difference a D.O. makes: Osteopathic medicine in the twentieth century*. Oklahoma City: Times-Journal Publishing Company.

Kane, R.L. (1997). *Understanding health care outcomes research.* Maryland: Aspen Publications.

Lederman, R.P. (1992). Exploratory data analysis part I: Skepticism and openness. *MCN, 17*, 325.

Lederman, R.P. (1993). Exploratory data analysis part II: Steps in examining data. *MCN, 18*, 47.

Licciardone, J.C., Stoll, S.T., Fulda, K.G., Russo, D.P., Siu. J., Winn, W. & Swift, J., Jr. (2003). Osteopathic manipulative treatment for chronic low back pain: A randomized controlled trial. *Spine, 28*(13), 1355-1362.

Meyer, P. (2003). *Osteopathic manipulative treatment and Carpal Tunnel Syndrome: A pilot study (Research proposal)*. Unpublished manuscript, University of North Texas Health Science Center, Fort Worth.

Noll, D.R., Shores, J.H., Gamber, R.G. & Slocum, P.C. (1998). The efficacy of adjunctive OMT in the elderly hospitalized with pneumonia (Abstract). *The Journal of the American Osteopathic Association, 98*(7), 389.

Noll, D.R., Shores, J.H., Gamber, R.G., Herron, K.M. & Swift, J., Jr. (2000). Benefits of osteopathic manipulative treatment for hospitalized elderly patient with pneumonia. *The Journal of the American Osteopathic Association, 100*(12), 776-781.

Neutens, J.J. & Rubinson, L. (2002). *Research techniques for the health sciences, 3rd ed.* San Francisco: Benjamin Cummings.

Punch, K.F. (2000). *Developing effective research proposals*. London: Sage Publications.

Scandalis,T.A., Hallas, B.H., Wells, M.R. & Bosak A. (1997). The effect of OMT on range of motion in the arthritic knee (Abstract). *The Journal of the American Osteopathic Association, 97*(8), 484.

Shelly, M.A. (1996). Exploratory data analysis: Data visualization or torture?. *Infection Control and Hospital Epidemiology, 17*(9), 605-612.

Smutny, C.J., Wells, W.R. & Bosak, A. (1998). Osteopathic considerations in Parkinson's disease (Abstract). *The Journal of the American Osteopathic Association, 98*(7), 389.

Steele, K.M., Kukulka, G. & Ikner, C.L. (1997). Effect of osteopathic manipulative treatment (OMT) on childhood otitis media outcomes (Abstract). *The Journal of the American Osteopathic Association,97*(8), 484.

Tukey, J.W. (1977). *Exploratory data analysis*. Menlo Park: Addison Wesley.

Verran, J.A. & Ferketich, S.L. (1987). Exploratory data analysis: Comparisons of groups and variables. *Western Journal of Nursing Research, 9*(4), 617-625.

Verran, J.A. & Ferketich, S.L. (1987). Exploratory data analysis: Examining single distributions. *Western Journal of Nursing Research, 9*(1), 142-149.

Victor, N. (1982). Exploratory data analysis and clinical research. *Method of Information in Medicine, 21*, 53-54.

Witte, R.S. (1985). *Statistics (2nd ed.).* New York: Holt, Rinehart and Winston.