DTIC FILE COPY

(1)

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOC

Form Approved
OMB No. 0704-0188

AD-A217 907

1a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

NONE

2a. SECURITY CLASSIFICATION AUTHORITY

3. AVAILABILITY OF REPORT
APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED.

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFIT/CI/CIA-89-154

6a. NAME OF PERFORMING ORGANIZATION
AFIT STUDENT AT
VA Commonwealth UNIV

6b. OFFICE SYMBOL (If applicable)

7a. NAME OF MONITORING ORGANIZATION
AFIT/CIA

6c. ADDRESS (City, State, and ZIP Code)

7b. ADDRESS (City, State, and ZIP Code)
Wright-Patterson AFB OH 45433-6583

8a. NAME OF FUNDING / SPONSORING ORGANIZATION

8b. OFFICE SYMBOL (If applicable)

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State, and ZIP Code)

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification) (UNCLASSIFIED)
THE RELIABILITY OF CLINICAL MEASUREMENTS OF FORWARD BENDING OBTAINED BY THE USE OF THE MODIFIED FINGERTIP-TO-FLOOR METHOD

12. PERSONAL AUTHOR(S)
MICHAEL GEORGE GAUVIN

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
THESIS/~~DISSERTATION~~	FROM _____ TO _____	1989	108

16. SUPPLEMENTARY NOTATION
APPROVED FOR PUBLIC RELEASE IAW AFR 190-1
ERNEST A. HAYGOOD, 1st Lt, USAF
Executive Officer, Civilian Institution Programs

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

DTIC
ELECTE
FEB 12 1990
S D
B

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
ERNEST A. HAYGOOD, 1st Lt, USAF	(513) 255-2259	AFIT/CI

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

AFIT/CI "OVERPRINT"

The Reliability of Clinical Measurements of Forward Bending Obtained By Use Of The Modified Fingertip-to-Floor Method

by

Michael George Gauvin
B. S., University of Minnesota, 1981
M.S., Chapman College, 1985

Director: Jules M. Rothstein, Ph.D., P.T.
Associate Professor of Physical Therapy

Thesis submitted in partial fulfillment

of the requirements for the

Degree of Master of Science in the

Department of Physical Therapy

School of Allied Health Professions

Medical College of Virginia

Virginia Commonwealth University

Richmond, Virginia

August 18, 1989

90 02 12 012

This thesis by Michael G. Gauvin is accepted in its present form as satisfying the thesis requirement for the degree of Master of Science.

Date: Aug. 17, 1989

Approved: _____
Director of Thesis

Aug. 17, 1989 _____

Aug 17, 1989 _____

Approved: _____
Dean, School of Allied Health Professions

Acknowledgements

I thank my wife and sons for being patient and supportive while I worked on this thesis. I am grateful for the guidance given to me by Jules Rothstein and Dan Riddle, and also want to thank David Bauer for being there when I needed him.

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

Table of Contents

	Page
List of Tables	v
List of Figures	vi
Abstract	vii
Chapter I. Introduction	1
Statement of the Problem	7
Questions of the Study	7
Operational Definitions	8
Limitations	9
Assumptions	9
Chapter II. Literature Review	10
Clinical Methods for Measuring Lumbar Flexion	10
Instruments Placed Directly Over the Lumbosacral Area	10
The Fingertip-to-Floor Methods	26
Usefulness of FTF Measurements in Assessing Lumbar Flexion	29
Reliability of Measurements Obtained with the MFTF Method	33
Issues Related to Reliability	40
Studying the Reliability of Measurements	43
Statistical Tests Used to Determine Reliability	46
Summary	50
Chapter III. Research Method	52
Subjects	52
Testers	53

Recorder..... 54

Instrumentation..... 54

Method of Data Collection..... 54

Data Analysis..... 58

Chapter IV. Results..... 60

Intra-tester Reliability..... 60

Inter-tester Reliability..... 60

Chapter V. Discussion, Conclusions and Summary..... 62

Discussion..... 62

Suggestions for Further Study..... 68

Conclusions..... 68

Summary..... 69

References..... 71

Appendices..... 76

A. Subject Consent Form..... 76

B. Data Collection Form..... 78

C. Instructions to Therapists..... 82

D. Article..... 83

Vita..... 108

List of Tables

Table		Page
1	Intraclass Correlation Coefficients for Reliability.....	61
2	The Effect of Pain on the MFTF Distance.....	65

List of Figures

Figure		Page
1	Therapist Taking MFTF Measurement.....	56
2	Therapist Transferring the Tape Measure to the Recorder.....	57

Abstract

THE RELIABILITY OF CLINICAL MEASUREMENTS OF FORWARD BENDING OBTAINED USING THE MODIFIED FINGERTIP-TO-FLOOR METHOD

Michael G. Gauvin, Capt., BSC, USAF

Medical College of Virginia Campus, Virginia Commonwealth University, 1989

Major Director: Jules M. Rothstein, Ph.D., P.T.

The purpose of this study was to examine the intra and inter-tester reliability of measurements obtained using a modified version of the fingertip-to-floor (FTF) method of assessing forward bending. With the modified FTF (MFTF) method patients stand on a stool and forward bend so that measurements can be taken on patients who are able to reach beyond the floor. In this study randomly paired physical therapists took repeated MFTF measurements on 73 patients with low back pain. Intraclass correlation coefficients (ICC) were calculated for intra and inter-tester reliability. The ICC value for intra-tester reliability was 0.98 and the ICC value for inter-tester reliability was 0.95. The results of this study suggest that measurements of forward bending obtained on patients with low back pain using the MFTF method are highly reliable.

CHAPTER ONE

Introduction

Physical therapists who examine and treat patients with low back pain usually assess the patient's ability to flex the lumbar spine. Many therapists appear to assess the degree of limitation of lumbar flexion by use of observational techniques. These therapists do not actually measure the amount of flexion. However, some clinicians feel that it is important to measure the amount of lumbar flexion. Measurements of lumbar flexion have been used to assist in the diagnosis of conditions such as ankylosing spondylitis (Calabro, 1982; Macrae & Wright, 1969) and other forms of arthritis (Schober, 1937). Clinicians also use measurements of lumbar flexion to help determine whether there is a functional loss due to low back pain (Mayer, Tencer, Kristoferson & Mooney, 1984). In addition, the selection of treatment techniques may be partially based on the assessment of lumbar flexion (McKenzie, 1981; Mitchell, Moran & Pruzzo, 1979). Measurements of lumbar flexion are also used to assess patient progress (Cox, 1985; Maitland, 1986).

A variety of methods of measuring lumbar flexion have been described. Many physical therapists recommend visually estimating lumbar flexion (Buswell,

1982; Paris, 1979). Spondylometers (Twomey & Taylor, 1979; Taylor & Twomey, 1980), inclinometers (Loebl, 1967; Mayer et al., 1984) and standard goniometers (Cox, 1985) have been used to take angular measurements of lumbar flexion. Burton (1986) used a draftsman's flexible ruler to measure lumbar flexion. A tape measure has been used to take measurements over the lumbar spine and these measurements have been used to represent the amount of lumbar flexion (Macrae & Wright, 1969; Schober, 1937).

A commonly recommended method for assessing lumbar flexion of low back pain patients is the fingertip-to-floor (FTF) method (Hoppenfeld, 1976; Kendall & Jenkins, 1968; Kraus, 1970; Mitchell et al., 1979; Ponte, Jensen & Kent, 1984). Measurements are obtained with the FTF method by having the patient forward bend in an attempt to touch the floor with his fingertips. The patient is usually instructed to keep his knees extended during the measurement procedure. The distance between the tip of the middle finger and the floor is visually estimated or measured (Ponte et al., 1984).

Some authors have modified the FTF method by having the subject stand on a stool so that measurements could be taken on subjects who are able to

touch the floor with their fingertips or reach beyond the level of the floor (Broer & Galles, 1958; Buxton, 1957; Frost, Stuckey, Smalley & Dorman, 1982). Measurements are made with the modified FTF Method (MFTF) by having the subject forward bend in an attempt to touch the floor with his fingertips and then measuring the distance between the tip of the middle finger and the top of the stool.

Although the FTF method has been used to measure lumbar flexion the FTF distance is not just a measurement of flexion of the lumbar spine. When a patient bends forward in an attempt to touch the floor with his fingertips he can flex the hips and entire spine. He can also protract the scapulae, flex the shoulders, and extend the elbows, metacarpophalangeal and interphalangeal joints. Some authors argue that the FTF method should not be used to measure lumbar flexion because the FTF distance is dependent on motion occuring at a number of joints (Biering-Sorenson, 1984; Moll & Wright, 1987; Rae, Waddell & Venner, 1984). Individual differences in hip, spinal and upper extremity range of motion make the use of a single FTF measurement as a measurement of lumbar flexion questionable.

Serial measurements taken with the FTF method have been used to assess the efficacy of treatments used for low back pain (Doran & Newell, 1975; Godfrey, Moran & Schatzker, 1984; Haldeman, Gilles, Haldeman & Patterson, 1975; Kendall & Jenkins, 1968; Lidstrom & Zachrisson, 1970; Ponte et al., 1984). In these studies single measurements of the FTF distance were not used to make inferences on the actual amount of lumbar flexion of patients. Pre-treatment and post-treatment measurements were taken and the authors were interested in changes in the FTF distance. They apparently believed that changes in the FTF measurement were exclusively or primarily due to changes in the amount of lumbar flexion.

Ponte et al. (1984) compared the Williams (1955) exercise protocol with the McKenzie (1981) exercise protocol. The authors concluded that lumbar flexion increased more in patients treated with the McKenzie protocol. This conclusion was made because there was a greater decrease in the FTF distance of patients in the McKenzie group. Therapists who use the FTF method to measure changes in the lumbar flexion of patients with low back pain are essentially making assumptions. They are assuming that any change in the FTF distance is a result of a change in lumbar flexion because they are

assuming that motion in the hips, thoracic spine and upper extremities remains relatively unchanged from measurement to measurement.

Clinicians may not be correct in assuming that changes in the FTF distance reflect changes in the lumbar flexion of patients with low back pain. Correlating changes in the FTF distance with changes in measurements of lumbar flexion obtained from radiographs of the lumbar spine would reveal if the FTF method is a valid method for monitoring changes in lumbar flexion. However, such a study would require that patients be exposed to radiation at various intervals as they recovered from low back pain. Patients should not be exposed to radiation unnecessarily and radiographic techniques used for research are costly and time consuming, requiring special equipment such as stabilizing frames (Portek, Pearcy, Reader & Mowat, 1983). Therefore, before attempting a validity study a reliability study seems logical. The MFTF method would be the most appropriate method to use in a reliability study because with the MFTF method subjects who cannot reach the floor, as well as subjects who can reach the floor or beyond the floor, can be measured.

The reliability of MFTF measurements taken on patients with low back pain has not been studied. Buxton (1957), Broer and Galles (1958) and Frost et al. (1982) concluded that measurements made with the MFTF method were reliable in healthy subjects. Buxton (1957) and Broer and Galles (1958) were interested in using the MFTF method on healthy subjects and chose a sample from the population they wanted to study.

Frost et al. studied the reliability of measurements made with the MFTF method because they said they were interested in using the MFTF method to measure patients. However, in selecting their sample they excluded subjects with a history of back surgery and subjects who had visited a physician because of low back pain within the five year period prior to their study. There are characteristics unique to a population that can affect reliability. Therefore, a sample from the population of interest should be selected when studying the reliability of measurements. For example, patients with low back pain might experience increased pain when asked to forward bend. Changes in the amount of pain might affect the consistency of a measurement. Normal subjects typically have no pain during forward bending. Therefore, the reliability of measurements made with

the MFTF method on patients with low back pain might be different than the reliability of measurements taken on normal subjects. The function of a sample is to provide information that allows one to generalize to the population from which the sample was selected (Ghiselli, Campbell & Zedeck, 1981). The reliability of measurements may differ between populations (Rothstein, 1985). Therefore, the reliability of MFTF measurements taken on patients with low back pain needs to be assessed if clinicians are to use the MFTF method to measure patients with low back pain.

Statement of the Problem

Clinicians often use changes in the FTF distance to make inferences about changes in lumbar flexion. The MFTF method is simple and requires only a tape measure and a stool. However, the clinical reliability of MFTF measurements has not been determined. The purpose of this study was to assess the intra-tester and inter-tester reliability of measurements taken with the MFTF method on patients with low back pain.

Questions of the Study

The questions of this study were:

1. What is the intra-tester reliability of

measurements obtained with the MFTF method on patients with low back pain?

2. What is the inter-tester reliability of measurements obtained with the MFTF method on patients with low back pain?

Operational Definitions

Appropriate patient. An appropriate patient was a patient referred to the physical therapy clinic for treatment of low back pain who, in the opinion of the referring therapist, had a problem that required assessment of lumbar flexion. Patients had to be at least 18 years of age to be included in the study.

Forward bending. Forward bending was the movement an individual performed from the standing position by allowing the hips and spine to flex while maintaining full knee extension.

Recorder. The recorder was the primary author.

Referring therapist. The referring therapist was the therapist who identified an appropriate patient from his patient load. The referring therapist took the first set of measurements on the patient.

Re-test therapist. The re-test therapist took the second set of measurements on the patient.

Limitations

1. The sample used in this study was a sample of convenience.
2. The study was limited to therapists in one clinic so results may not be generalizable to therapists in other clinics.
3. Patients in this study were required to stand on a stool during the measurement. The results of this study may not be generalizable to measurements made with versions of the FTF method that do not require the patient to stand on a stool.

Assumptions

1. Requiring the therapists to use the blank side of the measuring tape to measure the MFTF distance did not affect the therapists' method of making the measurement.
2. There was no change in the MFTF distance as a result of repeated performance of forward bending.

CHAPTER II

Literature Review

The first section of this literature review will examine the clinical methods that have been described for measuring lumbar flexion. Most of these methods depend on placing some instrument over the lumbosacral area. These methods will be discussed first. Fingertip-to-floor methods will be discussed next. The last section will address general issues related to reliability.

Clinical Methods for Measuring Lumbar Flexion

Instruments placed directly over the lumbosacral area.

Twomey and Taylor (1979) described a lumbar spondylometer used for measuring lumbar flexion. The spondylometer consisted of two metal rods connected by a pivot joint. The free end of one rod was connected to a protractor which was placed over the sacrum. The free end of the second rod was placed at the level of the first lumbar vertebra (L1). When the subject flexed his spine the angle the caudal rod formed with the protractor was the angle of lumbar flexion. The design of Twomey and Taylor's spondylometer was based on the spondylometer described by Dunham (1949). Dunham's spondylometer was larger and was used for

measuring flexion of the entire thoracolumbar spine.

Taylor and Twomey (1980) assessed the reliability of measurements obtained with the lumbar spondylometer by having two examiners measure 12 subjects. No description of the sample was provided. The authors reported that the maximum difference between measurements obtained by the two examiners was 5 degrees and concluded that measurements of lumbar flexion obtained with the spondylometer are reproducible. However, because the authors did not use any statistical test to assess reliability, any conclusions on the reliability of spondylometer measurements are limited to the subjects in their study. Statistical tests that result in a probabilistic statement should be used if one is interested in making inferences from a sample to a population.

The clinical usefulness of the lumbar spondylometer has not been demonstrated. No studies were found in which the lumbar spondylometer was used to measure the lumbar flexion of patients. The reliability of measurements obtained with the lumbar spondylometer has not been adequately assessed in normal subjects or patients.

Burton (1986) used a draftsman's flexible ruler to measure upper and lower lumbar flexion. Marks were placed over the spinous process of the second sacral (S2), fourth lumbar (L4) and twelfth thoracic (T12) vertebrae. For the purposes of this study S2 was assumed to be at the level of the lower aspects of the posterior superior iliac spines (PSIS). The L4 and T12 levels were located by palpating and counting up from S2. Once the marks were applied the subject sat in a chair and flexed his lumbar spine. The flexible ruler was then molded to the shape of the lumbosacral area and the S2, L4 and T12 points were marked on the ruler. The ruler was then removed and the shape of the curve was traced onto a sheet of paper. The points corresponding to S2, L4 and T12 were also marked on the paper.

A tangent was drawn on the flexion curve at the S2, L4 and T12 points. Tangents were drawn by placing a ruler on the curve with a calibration mark at the points corresponding to the spinous processes. Burton did not report which point on the ruler was used for the calibration mark. The ruler was positioned so that points on the ruler 0.5 cm to either side of the calibration mark were equidistant from the curve. Burton did not report whether the distance between the

0.5 cm marks was measured or visually estimated. The angles formed by the intersection of tangents was measured with a protractor. The angle formed by the intersection of the L4 and T12 tangents was designated as the upper lumbar flexion angle. The angle formed by the intersection of the L4 and S2 tangents was designated as the lower lumbar flexion angle.

Burton assessed the intra-tester reliability of measurements obtained with the flexible ruler by having a single examiner take repeated measurements on 15 subjects. Burton did not provide a description of the sample. The Pearson product-moment correlation coefficient was used to assess the degree of reliability. The coefficient for upper lumbar flexion was 0.95 and the coefficient for lower lumbar flexion was 0.97. The high r values indicate a strong linear relationship but do not provide information on the degree of agreement between measurements. The reliability of measurements of lumbar flexion obtained with a flexible ruler should be studied further if the flexible ruler is to be used to measure the lumbar flexion of patients with low back pain. The flexible ruler does not appear to be in widespread clinical use.

Some clinicians have used a tape measure to take s..in distraction measurements over the lumbosacral area

(Macrae & Wright, 1969; Moll & Wright, 1971; Schober, 1937). These measurements have been used to reflect the extent of lumbar flexion. Schober was the first to describe a skin distraction method for measuring lumbar flexion. Schober placed a mark midway between the level of the posterior superior iliac spines (PSIS) and a second mark 10 cm cephalad to the PSIS mark. The patient flexed the lumbar spine and the new distance between the two marks was measured. The difference between the original measurement and the measurement obtained when the spine was flexed reflected the amount of lumbar flexion.

Macrae and Wright (1969) altered Schober's (1937) method by placing an additional mark 5 cm caudal from the PSIS mark. The authors stated that during lumbar flexion, the 5 cm mark moved considerably less than the PSIS mark, therefore, it could be used as a relatively fixed point from which to measure. Macrae and Wright measured the distance between the 5 cm mark and the 10 cm mark after the subject flexed the spine. The difference between the original 15 cm distance and the new distance was the amount of lumbar flexion.

Macrae and Wright (1969) measured 342 subjects using both versions of the distraction method. The authors examined the relationship between distraction

measurements and sex, age and clinical condition. The sample consisted of an unspecified number of patients with ulcerative colitis and their family members. No further description of the sample was provided. Eleven of the subjects participated in a validity study of the distraction measurements. Measurements obtained with Schober's (1937) method and with the modified Schober method were correlated with measurements of lumbar flexion obtained from radiographs. Radiographs were taken with the subjects standing upright and standing with their spines fully flexed.

Measurements were obtained from both radiographs by measuring the angle formed by two lines drawn on the radiographs. One line was drawn from the anterosuperior corner of the first lumbar vertebra. The second line connected the sacral promontory and a "convenient" bony landmark on the sacrum. The authors did not state if the same landmark was used for all 11 subjects. The difference between the measurements taken on the two radiographs was the angle of lumbar flexion. The authors reported that they superimposed one radiograph on the other in an attempt to ensure that the same landmarks were used on both radiographs. The reliability of measurements obtained from the radiographs was not reported.

Macrae and Wright (1969) assessed validity by calculating Pearson product-moment correlation coefficients. The coefficient for measurements obtained with the Schober method was 0.90. The coefficient for measurements obtained with the modified Schober method was 0.97.

Because of the apparent validity of distraction measurements some clinicians have used the modified Schober method for measuring lumbar flexion of patients with low back pain (Davies, Gibson & Tester, 1979; Evans, Burke, Lloyd, Roberts & Roberts, 1978). Other clinicians have considered the modified Schober method the most appropriate method for measuring lumbar flexion of patients with ankylosing spondylitis. (Moll & Wright, 1973; Rae et al., 1984). The modified Schober method has also been used for measuring normal subjects (Moll & Wright, 1971; Moran, Hall, Barr & Ansell, 1979). The purpose of these studies was to examine the changes in lumbar flexion with age and to examine the difference in lumbar flexion of normal male and female subjects.

Some clinicians have used inclinometers, also known as gravity dependent goniometers, to measure lumbar flexion (Burdett, Brown & Fall, 1986; Mayer, Tencer, Kristoferson & Mooney, 1984; Portek, Pearcy,

Reader & Mowat, 1983). An inclinometer consists of a fluid filled circular dial with a weighted needle. The dial is marked in one degree increments and is attached to a base to allow for placement over the spine. When the subject flexes the spine inclinometer measurements are taken at the thoracolumbar junction (TL) and at the level of the first sacral vertebra (S1). The angle of lumbar flexion is determined by subtracting the S1 measurement from the TL measurement.

Portek et al. (1983) assessed the intra-tester and inter-tester reliability for measurements of lumbar flexion obtained with an inclinometer and measurements obtained with the modified distraction method. The authors also assessed the reliability of measurements of lumbar flexion obtained from radiographs.

Portek et al. (1983) addressed intra-tester reliability by having one examiner take 10 measurements on a single subject with each of the methods studied. Measurements were used to calculate coefficients of variation (CV). The CV's were 16.4% for inclinometer measurements, 8.5% for distraction measurements and 4.0% for measurements obtained from the radiograph. The CV is the standard deviation expressed as a percentage of the mean (Steele & Torrie, 1960).

The results reported by Portek et al. (1983) indicate that the measurements obtained from the radiographs were more reliable than the inclinometer measurements and skin distraction measurements. However, Portek et al. (1983) did not adequately assess intra-tester reliability because they only studied the variability within one subject. The authors could have more adequately assessed reliability by taking repeated measurements on a large number of subjects representing a specific population. If the authors had taken measurements on a large number of subjects they would have had to use a statistical test other than the CV to assess reliability. The CV does not distinguish between true variation due to individual differences between subjects and variation due to measurement error.

Portek et al. (1983) assessed inter-tester reliability of measurements of lumbar flexion by having two examiners take measurements on 14 subjects with the three methods studied. The subjects were patients admitted for orthopedic surgery, who did not have low back pain. Portek et al. stated that measurements of lumbar flexion are taken on patients with low back pain and patients with rheumatic disease. However, in their assessment of reliability they exluded patients with

low back pain. The authors should have selected a sample from the population of interest instead of selecting patients representing a population, or populations, that normally do not have measurements of lumbar flexion taken. The authors used a paired t-test to determine inter-tester reliability. However, the paired t-test only reveals if there is a statistically significant difference between sets of measurements (Ott, 1984) and does not assess the degree of agreement or covariance between measurements.

Portek et al. (1983) attempted to assess the validity of inclinometer measurements and distraction measurements. The authors correlated measurements obtained with the clinical methods with measurements obtained from radiographs and used the Pearson product-moment correlation coefficient to determine validity. The coefficient for inclinometer measurements was 0.41 and the coefficient for distraction measurements was 0.43. The low coefficients revealed that there was not a strong linear relationship between the clinical measurements and the measurements obtained from radiographs. The authors concluded that inclinometer and distraction measurements are not valid. This conclusion is in contrast to Macrae and Wright's (1969) conclusion about the validity of distraction

measurements. However, Portek et al's. method of assessing validity leads one to question their study.

Portek et al. (1983) positioned the subjects differently for the three measurements of lumbar flexion. The radiographs used to obtain measurements of lumbar flexion were taken with the subject standing in a stabilizing frame with his pelvis stabilized and his arms resting on a platform at shoulder height. The inclinometer measurements were taken with the subject sitting and the distraction measurements were taken with the subject standing outside of the stabilization frame. The actual amount of lumbar flexion of the subjects may have varied due to the different positions used while obtaining measurements. Macrae and Wright (1969) positioned their subjects in a similar manner for the radiographs and the distraction measurements.

Burdett et al. (1986) assessed the reliability of measurements obtained with five methods of measuring lumbar flexion. Measurements were taken with an inclinometer, the distraction method, a parallelogram goniometer, and two "platform" methods. All of these methods, with the exception of the distraction method, were also used to take measurements of the lumbar lordosis and lumbar extension. The authors assessed inter-tester reliability by having two physical

therapists measure 23 subjects with each method. Intra-tester reliability was not studied. Subjects were between 20 and 40 years of age and had no history of "chronic low back pain" and were not "grossly overweight." The authors did not report the mean age or standard deviation and did not define "chronic low back pain" or "grossly overweight."

The subjects were positioned prone prior to being measured. Marks were placed over the midline of the trunk at the level of the PSIS, 10 cm cephalad from the PSIS mark and 3 cm caudal from the PSIS mark. These marks were used to take distraction measurements. A mark was also placed at the thoracolumbar junction. The subjects then stood and two 10 cm long wooden pointers mounted on platforms were attached to the midline of the trunk at the thoracolumbar junction and the level of their PSIS.

Measurements of the lumbar lordosis were taken while the subject was standing. The subject then sat on a chair with his thighs parallel to the floor, his knees at 90 degrees and his feet flat on the floor. Subjects were instructed to flex their trunks as far as possible or until their shoulders contacted their thighs. When the subjects attained full flexion they were instructed to place their forearms under their

thighs and grasp their opposite elbows with their hands. Once the subject was in the flexed position the angle between the wooden pointers was measured with a standard goniometer and a photograph was taken of the pointers. The standard goniometer was used later to measure the angle between the pointers on the photograph. The pointers were removed and the tape measure was used to measure the new distance between the 3 cm and 10 cm marks. Inclinometer measurements were made over the thoracolumbar junction and just caudal from the PSIS mark.

The last instrument used was the parallelogram goniometer. The goniometer consisted of two protractors connected by a plastic ruler hinged in the middle. The protractors were placed at the thoracolumbar junction and just caudal from the PSIS mark. The difference between the angles on the two protractors was the angle of flexion. When the measurements of flexion were completed the subject was positioned prone and the extension measurements were taken.

Burdett et al. (1986) determined inter-tester reliability for the measurements of lumbar flexion obtained with the five methods by calculating intraclass correlation coefficients (ICC). They

considered an ICC of 0.80 or greater to be an indication of high reliability. Burdett et al. did not report which form of ICC they used. Shrout and Fleiss (1979) described six forms of the ICC. Each form of ICC is appropriate for a specific purpose. Authors using the ICC should report which form they used and explain the rationale for their choice.

The ICC's reported by Burdett et al. (1986) were 0.91 for the goniometric measurements of the pointers, 0.92 for the measurements obtained from the photographs of the pointers, 0.72 for the distraction measurements, 0.85 for the inclinometer measurements and 0.87 for the parallelogram goniometer measurements. The authors concluded that measurements obtained with all the methods except the distraction method were reliable.

The clinical usefulness of Burdett et al's. (1986) study is limited because the authors measured normal subjects. The reliability of measurements taken on patients with low back pain, or some other condition, may differ from the reliability of measurements taken on normal subjects. The authors' method of studying reliability is also questionable. The first therapist took 13 measurements on each subject before the second therapist took any measurements. During the measurement session the subject had to maintain a flexed position while five measurements were taken and

to maintain an extended position while four measurements of extension were taken, The order of the measurements was not randomized. Fatigue, discomfort and changes in motivation may have affected the subjects' performance during a long measurement session. These factors could have resulted in individual changes in the subjects' spinal range of motion that may have been attributed to measurement error. Burdett et al. should have studied each method independently so that the time interval between paired measurements was shorter and the chance of individual variations in spinal range of motion was minimized.

Burdett et al. (1986) also attempted to assess the validity of the measurements they obtained. The authors correlated measurements obtained with the clinical methods with measurements obtained from radiographs. The validity of distraction measurements was not studied. The authors stated that they did not include the distraction method because two radiographs are required to study the validity of distraction measurements and they did not want to expose subjects to radiation more than one time. Measurements were taken on six normal subjects. Information about the subjects' ages was not provided. The authors calculated Pearson product-moment correlation

coefficients to assess validity. The coefficients ranged from 0.46 to 0.76. The authors concluded that none of the measurements were valid because all of the coefficients were less than 0.80.

Burdett et al's. (1986) method of assessing validity is questionable. Subjects were sitting when the clinical measurements were taken but were standing when the radiographs were taken. The amount of lumbar flexion in the sitting position may differ from the amount of flexion when standing. When a subject flexes the spine in a sitting position he may not be able to flex as far as when standing because of contact between the chest and thighs. Individual differences in subjects' thigh mass and chest circumference would probably result in varying differences between the amount of sitting and standing flexion. Therefore, it is possible that there would not be a linear relationship between sitting and standing flexion measurements. Burdett et al's. results might have been different if they had positioned subjects in a similar manner for all measurements.

A review of the literature has revealed that a number of instruments are placed over the lumbosacral area to obtain measurements of lumbar flexion. With the exception of skin distraction measurements,

measurements obtained with these instruments have not been shown to be valid. Further study on the validity of measurements obtained with these methods should be undertaken to determine if clinicians are to be justified in making inferences about lumbar flexion based on these measurements. Validity of distraction measurements has only been demonstrated in a single study (Macrae & Wright, 1969). Further study on the validity of distraction measurements would strengthen the argument for the use of the distraction methods.

The reliability of measurements of lumbar flexion also needs to be studied. Therapists use measurements to monitor patient progress. If the reliability of a measurement is unknown then the therapist cannot confidently attribute a change in the measurement to a true change in the variable being measured because the change could also be due to measurement error.

The fingertip-to-floor methods

Some clinicians have used the fingertip-to-floor (FTF) method for assessing lumbar flexion (Hoppenfeld, 1976; Mitchell et al., 1979; Taylor, 1983). The FTF method differs from the methods previously discussed because FTF measurements are not taken directly over the lumbosacral area. FTF measurements are obtained by having the patient forward bend in an attempt to touch

the floor with his fingertips. The distance between the tip of the middle finger and the floor is usually estimated or measured with a tape measure (Ponte et al., 1984) or yardstick (Broer & Galles, 1958).

One obvious problem with the FTF method is that measurements cannot be taken on patients who are able to touch the floor with their fingertips. Biering-Sorenson (1984) correlated FTF measurements with measurements taken with Macrae and Wright's (1969) skin distraction method. Measurements were taken on 449 male subjects and 479 female subjects. The ages of the subjects spanned the fourth through seventh decades. Biering-Sorenson used the Pearson product-moment correlation coefficient to assess the degree of correlation between measurements. The author reported a coefficient of -0.35 for males and a coefficient of 0.22 for females.

The results of this study revealed that there was not a linear relationship between FTF measurements and distraction measurements. However, Biering-Sorenson's (1984) method of measuring the FTF distance did not allow for appropriate measurements of subjects who were able to touch the floor with their fingertips. The author stated that if a subject was able to touch the floor he was not measured but assigned a value of zero.

Therefore, all subjects who were able to touch the floor had identical FTF measurements while their distraction measurements may have varied. If Biering-Sorenson had been able to actually take measurements on these subjects, instead of just assigning them values of zero, the relationship found between FTF measurements and distraction measurements may have been different.

Biering-Sorenson (1984) could have taken measurements on all subjects by using a modified FTF (MFTF) method that has been described (Broer & Galles, 1958; Buxton, 1957; Frost et al., 1982). With the MFTF method the subject stands on a stool, bends forward and reaches toward the floor with his fingertips. The distance between the subject's fingertips and the top of the stool is measured. If the subject reaches the top of the stool a zero is recorded. Positive and negative values are used for measurements above and below the top of the stool. Frost et al. (1982) recorded measurements above the top of the stool as negative values. Buxton (1957) and Broer and Galles (1958) used negative values for measurements below the top of the stool.

Kipper and Parker (1987) correlated MFTF measurements and measurements of thoracolumbar flexion

obtained from photographs. Measurements were taken on 16 healthy men and 17 healthy women. The mean age of the subjects was 21.6 years with a standard deviation of 3.4 years. Out of the 33 subjects, 20 were able to reach beyond their toes, and the mean MFTF measurement was -3.0 cm with a standard deviation of 8 cm. If Kipper and Parker had not had their subjects stand on a stool they would have been able to take measurements on only 13 of the 33 subjects.

The MFTF method not only allows for measurements from subjects who are able to reach the floor, but also allows measurements to be taken that can reflect changes in patient status. Physical therapists use FTF measurements to monitor patient progress. With the MFTF method the therapist can monitor changes in patients who are able to reach the floor.

Usefulness of FTF measurements in assessing lumbar flexion

Some clinicians have stated that individuals with "normal" lumbar and hip motion should be able to forward bend and touch the floor with their fingertips (Kraus, 1965, 1970; Mitchell et al., 1979). Other clinicians argue that the FTF distance should not be used to make a determination of whether a patient has "normal" lumbar flexion because the ability of a

patient to touch the floor is dependent on motion at the hips, spinal joints and joints of the upper extremities (Finneson, 1980; Kubinec, 1977; Rae et al., 1984).

Rae et al. (1984) reported that some patients with ankylosing spondylitis were able to touch the floor with their fingertips even though they were unable to flex their lumbar spines. Finneson (1980) reported similar findings in patients who had their lumbar spines surgically fused. Biering-Sorenson (1984) and Kipper and Parker (1987) reported that single FTF measurements did not correlate highly with single measurements of spinal flexion. Because of individual differences in hip, spinal and upper extremity range of motion it is unlikely that two individuals with the same FTF distance would have the same amount of lumbar flexion.

Biering-Sorenson (1984) compared single FTF measurements with single measurements of lumbar flexion taken with the distraction method (Macrae & Wright, 1969). Measurements were taken on 479 female subjects and 449 male subjects. The author reported correlation coefficients of $\underline{r} = -0.35$ for males and $\underline{r} = 0.22$ for females.

Kipper and Parker (1987) compared single MFTF measurements with single measurements of thoracolumbar flexion. The measurements were obtained from photographs of 33 subjects. The angle of thoracolumbar flexion was the angle between a line connecting points representing the first thoracic vertebra and the posterior superior iliac spine (PSIS) and another line connecting the PSIS point and a point representing the anterior superior iliac spine.

Kipper and Parker used the Pearson product moment correlation coefficient to compare MFTF measurements with measurements of lumbar flexion and reported a coefficient of 0.1. The results revealed that there was not a linear relationship between the two sets of measurements. However, the usefulness of Kipper and Parker's study is limited because the authors did not assess the reliability or validity of their measurements of thoracolumbar flexion.

Biering-Sorenson (1984) and Kipper and Parker seem to have given support to the argument that single FTF or MFTF measurements are not representative of motion occurring in the spine. However, these authors did not provide information on the relationship between changes in the FTF distance and changes in lumbar flexion.

Patients with low back pain commonly are able to flex the lumbar spine more as they improve. An increase in lumbar flexion results in a decrease in the FTF distance because the ability of a patient to reach toward the floor is partially dependent on motion occurring in the lumbar spine. The relationship between changes in lumbar flexion and changes in the FTF distance has not been determined.

Many clinicians have used FTF measurements to monitor changes in the amount of lumbar flexion (Doran & Newell, 1975; Godfrey et al., 1984; Haldeman et al., 1975; Kendall & Jenkins, 1968; Lidstromm & Zachrisson, 1970; Ponte et al., 1984). These clinicians obtained serial FTF measurements as patients recovered from low back pain and attributed any change in the FTF distance to a change in the amount of lumbar flexion.

Validity refers to the appropriateness of inferences made from measurements (American Psychological Association, 1974). Clinicians who use the FTF distance to document changes in lumbar flexion infer that a change in the FTF distance corresponds with a change in lumbar flexion. However, the validity of serial FTF measurements has not been determined. When a measurement is used as a substitute for the characteristic of interest, in this case the change in

lumbar flexion, the validity of the measurement should be determined (Ghiselli, Campbell & Zedeck, 1981).

The validity of serial measurements of the FTF distance could be determined by correlating changes in the FTF distance with changes in measurements of lumbar flexion obtained from radiographs. Serial measurements would be taken as patients recovered from low back pain, requiring patients to be exposed to radiation on a number of occasions. Because patients should not be exposed to radiation unnecessarily and radiographic procedures are time consuming and costly, practicality dictates that the reliability of measurements of the FTF distance should be determined prior to a study of validity.

Reliability of measurements obtained with the MFTF method

Buxton (1957) assessed the reliability of measurements obtained with the MFTF method by taking repeated measurements of the FTF distance of 50 children ranging in age from six to 15 years. The mean and standard deviation of the childrens' ages were not reported. No further description of the sample was provided. Subjects were instructed to stand on a 6 inch high stool and bend forward in an attempt to touch the floor with their fingertips. The distance between

the fingertips and the toes was measured with a yardstick to the nearest half-inch. Buxton did not report which fingertip was used to measure from, or which point on the toes were used to obtain the measurement. If the subject reached his toes a zero was recorded. If the subject was not able to reach his toes a negative value was recorded and if the subject reached beyond his toes a positive value was recorded. Measurements were taken during two sessions with two measurements taken per session.

Buxton (1957) assessed inter-session reliability by comparing the first measurement in session one with the first measurement in session two and by comparing the second measurement in session one with the second measurement in session two. Buxton reported a reliability coefficient of 0.95 for the first measurements and a coefficient of 0.96 for the second measurement. The author did not report what reliability coefficient was used. Buxton also did not report the number of testers who took measurements so it cannot be determined whether intra-tester or inter-tester reliability was studied.

Buxton's (1957) results reveal that MFTF measurements obtained on healthy school children may be reliable. However, the usefulness of this study to the

clinician is limited because the sample did not consist of subjects representing a patient population.

Broer and Galles (1958) assessed the intra-tester reliability of MFTF measurements by having an examiner take repeated measurements on 50 female university students. The women were 18 to 31 years of age and were all enrolled in physical education classes at the time of the study. Means and standard deviations for the subjects' ages were not reported. The subjects reached toward the floor while standing on a 14 inch high stool. The examiner used a yardstick to measure the distance between the fingertips and the top of the stool. The distance was measured to the nearest half-inch. Measurements above the stool were recorded as negative values and measurements below the stool were recorded as positive values. If the subject's fingertips were at the level of the top of the stool a zero was recorded.

Broer and Galles (1958) used the Pearson product-moment correlation coefficient (r) to assess reliability and reported a coefficient of 0.97. The authors concluded that measurements obtained with the MFTF method were reliable. The results of this study reveal that MFTF measurements may be reliable in healthy young women. However, a high r value can be

obtained if paired measurements consistently covary regardless of the difference between measurements. Although there was a strong linear relationship between measurements, the degree of agreement between measurements is not evident based on the information reported by the authors. The usefulness of this study to clinicians is limited because, like Buxton (1957), Broer and Galles did not measure patients.

Frost et al. (1982) assessed the reliability of measurements of forward bending, backward bending, side bending, rotation, straight-leg-raising, and prone knee flexion. The authors used the MFTF method to measure forward bending. Subjects in the study were 12 males and 12 females ranging in age from 20 to 55 years with a mean age of 33.8 years. The standard deviation of the age of the sample was not reported. The subjects had no history of back pain for the past five years and no history of back surgery. The examiners in the study were three physical therapists with 1.5 to 6.5 years of experience. The therapists were required to learn specific measurement protocols for each motion measured.

Measurements were taken during two measurement sessions. There was a one week time interval between sessions. Subjects performed "stretching" exercises

prior to each session. These exercises consisted of five repetitions of each motion that was measured. Therapists took three consecutive measurements of each motion. The sequence of the motions was forward bending, backward bending, right side bending, right rotation, right straight-leg-raising and right prone knee flexion. A therapist took all 18 measurements before the next therapist took any measurements on the subject. Each therapist measured all 24 subjects.

Frost et al. (1982), like Burdett et al. (1986), had one therapist take many measurements on a subject before the next therapist took any measurements on the subject. Therefore, there may have been individual changes in spinal and hip range of motion. As in Burdett et al.'s study, fatigue, discomfort and changes in motivation may have affected a subject's performance over the course of a long measurement session. Frost et al. should have studied each motion individually and done a number of studies instead of assessing the reliability of six different measurements in one study.

Frost et al. (1982) measured the FTF distance with the MFTF method. Subjects stood on a stool with their heels together, their knees "straight" and their arms in "neutral." The instructions given to the subjects were "bend forward as far as you can, keep your knees

straight". The distance between the tip of the right middle finger and the top of the stool was measured with a metal tape measure.

Frost et al. (1982) used one of the reliability coefficients described by Winer (1971) to assess reliability. Intra-tester and inter-tester reliability within sessions and between sessions was studied. The authors reported five coefficients ranging from 0.82 to 0.98 and concluded that MFTF measurements were reliable. However, the results of this study should be considered with caution because the authors did not report which one of the coefficients described by Winer they calculated. One of Winer's reliability coefficients is not sensitive to differences between measurements if measurements consistently covary (Bartko, 1976). This coefficient can be high even though the difference between measurements is large.

The results of Frost et al.'s (1982) study reveal that measurements obtained with the MFTF method may be reliable. However, because the authors did not take measurements on patients the usefulness of this study to clinicians is limited. When the reliability of a measurement is studied a sample should be selected from the population of interest. Frost et al. stated that they were interested in studying measurement methods

that could be used to quickly obtain accurate measurements with a minimum of patient discomfort. However, the authors purposely excluded patients with low back pain from participating in their study.

The usefulness of Frost et al.'s (1982) version of the MFTF method is questionable because they standardized subject positioning. In the clinical setting patients may not be able to assume a standardized position, such as standing with their heels together. Some patients may be more comfortable with their feet apart. Patients with low back pain often try to remove weight from the lower extremity on the involved side by leaning to one side. When studying the reliability of measurements the effects of individual differences in patient positioning and performance can be studied during a posteriori analyses.

A review of the literature has revealed that MFTF measurements may be reliable when taken on healthy subjects. However, the reliability of MFTF measurements taken on patients has not been assessed. The reliability of MFTF measurements taken on patients should be assessed to determine if the MFTF method is appropriate for use in the clinical setting.

Issues Related to Reliability

Physical therapists take repeated measurements on patients over time. Changes in these measurements are attributed to changes in the variable being measured. However, changes in measurements can also be partially due to measurement error. The amount of change due to error is dependent on the reliability of the measurement. Reliability has been defined as the consistency of a measurement (Bartko & Carpenter, 1976; Ghiselli, Campbell & Zedeck, 1981). If a measurement is highly reliable then therapists are probably correct in assuming that a change in the measurement reflects true variation. However, if a measurement has poor reliability the therapist must consider the possibility that a change in the measurement is due largely to measurement error.

There are a number of factors, or sources of error, that can threaten reliability. Errors made by the individual taking the measurements can result in inconsistent measurements (Rothstein, 1985). A single therapist taking repeated measurements may not perform the measurement procedure consistently. Many methods for measuring lumbar flexion require the therapist to palpate and mark points over the lumbosacral spine. Inconsistent palpation can result in inconsistent

placement of the device used to take measurements.

If two or more therapists take measurements on the same patient the chance of error is potentially greater because of individual differences in the methods used by the therapists. Therapists may have different methods for palpating spinous processes. The method of applying the measurement device may also differ. For example, when taking skin distraction measurements (Macrae & Wright, 1969) the examiner uses a tape measure to locate and mark a point 10 cm cephalad from a mark at the level of the PSIS. However, Macrae and Wright did not specify if the tape measure was pulled taut or if it was molded to the shape of the lumbar curve. Two therapists might apply the tape measure differently and obtain different measurements.

Expectation bias can also be a cause of measurement error. If a therapist is taking repeated measurements and is aware of the first measurement, the therapist will have an expectation of subsequent measurements. This expectation might influence the way in which the therapist takes subsequent measurements.

The examiner is not the only potential source of error. Changes in the patient being measured can also affect the consistency of a measurement (Rothstein, 1985). Many factors can affect the patient's

performance when MFTF measurements are taken. Patients with low back pain often experience an increase in pain during forward bending. Variations in pain might result in inconsistent performance of forward bending. Patients may also experience fear, anxiety, depression or fatigue. Variation in any of these factors could affect the reliability of MFTF measurements.

Flaws in the measurement device can be a third source of error (Rothstein, 1985). The instrument for obtaining MFTF measurements includes the tape measure, the footstool, the manner in which the examiner uses the tape measure and the instructions the examiner gives to the patient. Variation of the tape measure or footstool is unlikely. Flaws in the instrument used to take MFTF measurements would most likely be due to inconsistent device placement or inconsistent or confusing instructions. In the case of MFTF measurements flaws in the instrument are synonymous with errors made by the examiner.

The reliability of measurements can be affected by a number of factors. These factors need to be considered by clinicians who take measurements on a daily basis and by researchers, especially those who study the reliability of measurements.

Studying the reliability of measurements

The reliability of a measurement can be assessed with the test-retest method (Ghiselli et al., 1981). Repeated measurements are taken and the degree of agreement between repeated measurements is examined.

When studying the reliability of MFTF measurements two types of reliability can be examined. Intra-tester reliability refers to the ability of an individual to obtain consistent measurements (Rothstein, 1985) and is assessed by having a single examiner take repeated measurements. The degree of agreement between paired measurements is examined. If intra-tester reliability is poor the usefulness of the measurement is questionable because differences between repeated measurements are primarily due to measurement error and may not reflect a true change in the variable being measured.

Inter-tester reliability is the stability of a measurement between examiners (Rothstein, 1985) and is assessed by having two examiners take measurements on the same patient. High inter-tester reliability is necessary if two or more clinicians take measurements on a patient as the patient progresses. If inter-tester reliability is poor then any change in the measurement may be due to inconsistency between

therapists.

A measurement may have good intra-tester reliability but poor inter-tester reliability because of differences in the methods used by therapists. For example, two therapists might give different instructions to patients when taking MFTF measurements. One therapist might instruct the patient to bend forward as far as possible and the other therapist might instruct the patient to bend forward to the point of pain. The measurements obtained by the two therapists might differ greatly. If only intra-tester reliability is high the same therapist should take serial measurements on a patient.

One factor that needs to be considered when studying reliability is the time interval between repeated measurements. Frost et al. (1982) studied the reliability of MFTF measurements taken on normal subjects. The authors studied the reliability of repeated measurements taken on the same day and also studied the reliability of repeated measurements with a one week time interval between measurements. The authors reported good reliability regardless of the length of the interval.

The use of a one week time interval when studying the reliability of MFTF measurements taken on patients

with low back pain would be questionable. One might expect the FTF distance of normal subjects to remain constant from week to week. However, the FTF distance of patients with low back pain can change considerably after a single treatment session. Therefore, when studying the reliability of MFTF measurements taken on patients with low back pain a short time interval should be used in an attempt to assure that the variable of interest remains unchanged between measurements.

Expectation bias also needs to be considered when studying reliability. If the examiners record their own measurements, knowledge of the first measurement might influence subsequent measurements. Expectation bias can be controlled by preventing the examiner from knowing measurements at the time they are taken. Some researchers who have studied the reliability of goniometric measurements have covered the scales on the goniometers so that only the recorder could read the measurement (Riddle, Rothstein & Lamb, 1987; Rothstein, Miller & Roettger, 1983). In a study of the reliability of MFTF measurements expectation bias could be eliminated by covering one side of the tape measures and only allowing the recorder to see the side of the tape measure with the measurement scale.

Another issue that requires consideration when designing a reliability study is sample selection. The function of a sample is to provide information that allows the researcher to generalize to the population from which the sample was selected (Ghiselli et al., 1981). Frost et al. (1982) assessed the reliability of MFTF measurements taken on normal subjects. However, one of their reasons for studying reliability was to determine if the MFTF method was appropriate for use in the clinic. Because Frost et al. measured normal subjects generalization to a patient population is inappropriate. There are characteristics unique to a population that can affect reliability. Patients with low back pain often experience an increase in pain during forward bending. The amount of pain may vary. Some patients experience fear or anxiety during an examination. This fear or anxiety may vary as the patient adjusts to the clinical setting or as pain decreases. All of these factors can affect the consistency of MFTF measurements. These factors are not usually present in normal subjects.

Statistical tests used to determine reliability

Broer and Galles (1958) studied the reliability of measurements obtained with the MFTF method. The authors used the Pearson product-moment correlation

coefficient (r) as a reliability coefficient. The r is not an appropriate test to use when studying the degree of agreement between measurements. The r assesses covariance between measurements and if measurements consistently covary a high r will be obtained even if there is a large difference between paired measurements. For example, if two examiners obtained identical MFTF measurements the r will be 1.0. However, if there is a consistent difference of 5 cm. between paired measurements the r will also be 1.0. Even though the coefficients were identical the measurements in the first example were more reliable.

The r is an appropriate statistic to use for comparing measurements of two different variables. Macrae and Wright (1969) compared skin distraction measurements with measurements of lumbar flexion obtained from radiographs. The authors used the r to determine the degree of covariance between measurements and reported that there was a strong linear relationship between measurements. Based on their results, Macrae and Wright concluded that distraction measurements are valid. The r may not be an appropriate statistic for studying reliability but it can be used to assess the validity of a measurement.

Portek et al. (1983) used the paired t-test to assess inter-tester reliability of inclinometer and distraction measurements. The authors reported that there was no significant difference between sets of inclinometer measurements but there was a significant difference between sets of distraction measurements. The authors concluded that inclinometer measurements were more reproducible than distraction measurements. However, the paired t-test only assesses whether there is a significant difference between sets of measurements (Ott, 1984) and does not provide information on the degree of agreement between paired measurements.

Biering-Sorenson (1984) assessed the reliability of distraction measurements by taking repeated measurements on 127 subjects. Measurements were used to calculate a coefficient of variation (CV). The CV is the sample standard deviation expressed as a percentage of the sample mean (Steele & Torrie, 1960). Biering-Sorenson reported a CV of 4.83% and concluded that distraction measurements were reproducible. However, the CV does not adequately assess reliability because it cannot distinguish between measurement error and variation due to true differences between subjects. The standard deviation reflects the variability within

a sample. A large standard deviation can reflect true variability, variability due to error or a combination of both. Therefore, a measurement might be reliable even though the CV is large if the true variability between subjects is great.

Burdett et al. (1986) used a form of the intraclass correlation coefficient (ICC) to assess the reliability of measurements taken with a number of methods for measuring lumbar flexion. The authors did not report which form of ICC they used. Shrout and Fleiss (1979) described six forms of the ICC and each form is appropriate for a specific purpose. The ICC's are based on an analysis of variance and are appropriate for assessing reliability because they can be used to distinguish between true variation and variation due to error. If the error variance increases the ICC will decrease. An ICC of 1.0 can only be obtained if paired measurements are identical.

The selection of a form of ICC is dependent on the design of the study. One approach to studying the reliability of a measurement is to compare single measurements obtained by randomly paired testers. The most appropriate form of ICC to use when testers are randomly paired and single measurements are compared is ICC (1,1).

Summary

A variety of clinical methods for measuring lumbar flexion have been described. Measurements are obtained with most of these methods by placing some instrument over the lumbosacral area.

The FTF method is commonly recommended as a method for assessing lumbar flexion. However, because FTF measurements are dependent on motion occurring in the hips, spine and upper extremities the practice of using a single FTF measurement to make inferences on a patient's lumbar flexion is questionable. Some clinicians have used serial FTF measurements to monitor changes in lumbar flexion. These clinicians assume that as a patient improves, any change in the FTF distance is due to a change in lumbar flexion.

The MFTF method may be an appropriate method for documenting changes in a patient's lumbar flexion. The MFTF method allows the therapist to take measurements on patients who are able to touch the floor with their fingertips. MFTF measurements are quickly and easily obtained. The only items required are a tape measure and a stepstool. Both are readily available in most clinics. However, the reliability of MFTF measurements taken on patients with low back pain has not been

studied. The reliability of MFTF measurements taken on normal subjects has been studied but the results of these studies are of limited value to clinicians who measure patients with low back pain. The reliability of MFTF measurements taken on patients may differ from the reliability of measurements taken on normal subjects. There are characteristics unique to patients with low back pain that can affect reliability. Therefore, the reliability of MFTF measurements taken on patients with low back pain should be studied as a first step in determining the clinical usefulness of MFTF measurements.

CHAPTER THREE

Research Method

This study assessed the intra-tester and inter-tester reliability of measurements of forward bending obtained with the MFTF method on patients with low back pain. This chapter describes the subjects and testers who participated in the study, the instruments used in the study, and the methods used for collection and analysis of data.

Subjects

Subjects in this study were 73 patients referred for treatment of low back pain to the Department of Physical Therapy, Malcolm Grow Medical Center, Andrews Air Force Base, Maryland. Patients were included in the study if assessment of the amount of lumbar flexion was considered, by the referring therapist, to be an appropriate part of their examination. All patients included in this study were at least 18 years of age. Patients were asked to participate only if the referring therapist felt the patient could tolerate repeated forward bending. All patients were asked to read and sign a consent form (Appendix A) prior to inclusion in the study.

After the patient signed the consent form the patient was interviewed by the investigator and asked

questions pertaining to his back pain. Information obtained during the interview was recorded on the data collection form (Appendix B) to be available for use in a posteriori analyses to determine if age, sex, pain or diagnosis had any affect on reliability.

The subjects in this study consisted of 47 males and 26 females ranging in age from 18 to 73 years. The mean age of the subjects was 43.5 years with a standard deviation of 13.9 years.

Testers

Testers for this study were six physical therapists on the staff of the Department of Physical Therapy, Malcolm Grow Medical Center, Andrews Air Force Base, Maryland. All testers were given a description of the MFTF method (Appendix C) by the investigator. None of the therapists reported they routinely took FTF or MFTF measurements.

Prior to the beginning of the study each therapist was assigned a number. The investigator used these numbers and a table of random numbers (Ott, 1984) to make a random number list for each therapist. These lists were used to identify re-test therapists during the study.

Recorder

The recorder for this study was the primary investigator. Measurements were recorded on the data collection form (Appendix B).

Instrumentation

In this study a cloth tape measure* marked in .10 cm increments was used to obtain measurements. The reverse side of the tape measure was covered with silver duct tape to prevent examiner bias. The recorder checked the accuracy of the tape measure twice a day by comparing the centimeter scale of the tape measure with the centimeter scale on a metal meter stick. This was done to ensure that the length of the tape measure did not change with repeated use. The length of the tape measure did not change during the study.

Method of Data Collection

The recorder was notified when a therapist identified an appropriate patient for the study. The recorder described the study to the patient and asked the patient to read and sign the consent form. If the

---

*Cloth Tape Measure, Dritz Corporation, Spartanburg, SC, 29304

patient agreed to participate in the study the recorder interviewed the patient. When the interview was completed the re-test therapist was identified using the referring therapist's list of random numbers.

The referring therapist instructed the patient to stand on a 12-3/4 inch (32.4 cm) high stool and bend forward. Using the blank side of the tape measure the therapist measured the distance between the patient's right middle finger and the top of the stool. The therapist marked the distance on the tape measure with his thumbnail (see Figure 1). The therapist then held the tape measure up so that the recorder could see the marked side of the tape measure. The recorder used his thumbnail to mark the point on the centimeter scale that corresponded with the referring therapist's thumbnail and then took the tape from the referring therapist (see Figure 2). The referring therapist was not allowed to see the marked side of the tape measure and was not informed of the measurement.

The MFTF distance was recorded to the nearest .10 cm. A positive value was recorded if the patient was unable to reach the top of the stool and a negative value was recorded if the patient was able to reach beyond the top of the stool. A zero was recorded if the patient reached the top of the stool.

Figure 1. Therapist taking MFTF measurement.

Figure 2. **Therapist transferring the tape measure to the recorder.**

After the first measurement was taken the patient was allowed to step off the stool and move around the measurement area for 15 to 30 seconds. The patient then stepped back on the stool and the referring therapist took a second measurement. After the referring therapist took two measurements the re-test therapist was asked to report to the measurement area and take two measurements using the same method as the referring therapist used. The re-test therapist was not informed of the measurements obtained by the referring therapist.

During the measurement session the recorder noted the instructions given to the patients, the manner in which the therapist was positioned and the manner in which the therapist used the tape measure (Appendix B). Differences in technique between therapists were noted so that a posteriori analyses could be performed to study the effects, if any, of variations in technique on the reliability of measurements taken with the MFTF method.

Data Analysis

Intraclass correlation coefficients (ICC [1,1] Shrout & Fleiss, 1979) were calculated to assess intra-tester and inter-tester reliability. ICC (1,1) was chosen because it is the appropriate form of ICC to use

when assessing the reliability of single measurements taken by testers who are randomly paired from a group of testers. ICC (1,1) is also the most conservative form of ICC.

The intra-tester reliability of measurements obtained with the MFTF method was determined by comparing the first and second measurements made by the referring therapist and by the re-test therapist. There were 146 paired measurements.

The inter-tester reliability of measurements obtained with the MFTF method was determined by comparing the first measurement made by the referring therapist with the first measurement made by the re-test therapist. There were 73 paired measurements.

CHAPTER FOUR

Results

This chapter presents the results of the statistical analysis performed on data collected during this study.

Intra-tester Reliability

The ICC value for all sets of paired measurements was .98. This value was calculated by comparing the first and second measurements taken by the referring therapist and the first and second measurements taken by the re-test therapist.

Inter-tester Reliability

The ICC Value for inter-tester reliability was .95. This value was calculated by comparing the first measurement taken by the referring therapist with the first measurement taken by the re-test therapist (see Table 1).

Table 1

Intraclass Correlation Coefficients for Reliability

	N	ICC
Intra-tester Reliability	146	.98[a]
Inter-tester Reliability	73	.95[b]

[a] This ICC was calculated by comparing the first and second measurements of the referring and re-test therapist.

[b] This ICC was calculated by comparing the first measurement taken by the referring therapist with the first measurement taken by the re-test therapist.

CHAPTER FIVE

Discussion, Conclusions and Summary

This chapter includes a discussion of the results, suggestions for further study, and conclusions. The conclusions will be followed by a brief summary of the study.

Discussion

The results of this study were similar to those of Frost et al. (1982) who reported high intra-tester and inter-tester reliability for MFTF measurements taken on healthy subjects. However, there are a number of methodological differences between the two studies.

Frost et al. (1982) studied only healthy subjects. Therefore, the value of Frost et al.'s study is limited because the reliability of MFTF measurements obtained on healthy subjects might differ from the reliability of MFTF measurements obtained on patients. In the present study 73 patients with low back pain were measured because the population of interest was patients with low back pain.

Selecting a sample from the population of interest is important because each population has characteristics that might affect the reliability of a measurement. Patients with low back pain commonly have

increased pain when forward bending. Frost et al. (1982) were unable to examine the effect of pain on the reliability of MFTF measurements because they used normal subjects.

In the present study 22 of the 73 patients stated that they had an increase in pain during forward bending. To determine if an increase in pain had any effect on the reliability of measurements, a posteriori analyses were performed. Separate ICC's were calculated for repeated measurements taken on patients who had an increase in pain with forward bending and patients who had no change in pain with forward bending. The ICC for patients who had an increase in pain was .96 and the ICC for patients who had no change in pain was .94. The ICC's for inter-tester reliability were essentially the same for patients who had no change in pain during forward bending and patients who had an increase in pain.

An increase in pain during forward bending apparently did not influence reliability but it may have affected the patient's ability to forward bend. The results of an a posteriori t-test revealed that the mean of the first MFTF measurements (20.65 $\pm$ 14.5 cm) for patients who had an increase in pain when forward bending was significantly greater ($p<.05$) than the mean

of the first MFTF measurements (12.6 $\pm$ 12.9 cm) for patients who had no increase in pain when forward bending (see Table 2). The results of the *t*-test suggest that patients who report an increase in pain when forward bending may have a decrease in the ability to forward bend when compared to patients who do not report an increase.

Another difference between the present study and the study by Frost et al. (1982) was the instructions the therapists gave to the subjects. In the study by Frost et al. the therapists who took the measurements used a specific set of instructions. In the present study the therapists were allowed to use their own instructions. No pair of therapists gave the same instructions. One therapist gave very brief instructions, such as "bend forward as far as you can." Another therapist gave very specific instructions, telling the patient to "bend forward by rolling the spine, bending at the hips and keeping the knees straight."

Frost et al.'s (1982) instructions called for the subjects to stand with their heels together. The therapists in the present study allowed the patients to stand in the position that was most comfortable for the patient. Most patients stood with their feet

Table 2

The Effect of Pain on the MFTF Distance

	MFTF[a] Measurement (cm)	
	$\bar{x}$	SD
Increase in pain with FB[b] (n=22)	20.65	14.5
No change in pain with FB[b] (n=51)	12.6	12.9

[a] Modified Fingertip-to-Floor

[b] Forward Bending

approximately shoulder width apart. None of the 73 patients measured in this study stood with their heels together. The therapists were allowed to give their own instructions to the patients and patient positioning was not standardized in order to more closely replicate clinical procedure. The results of this study reveal that standardized instructions and patient positioning are not required for MFTF measurements to be highly reliable.

The results of this study indicate that the intra-tester and inter-tester reliability of MFTF measurements taken on patients with low back pain is very high. Data collected during this study also indicate that the MFTF method may be preferable to the FTF method for therapists who take serial measurements on patients. Out of the 73 patients measured in this study, 20 (27%) were able to reach the top of the stool or beyond. A therapist would need to use the MFTF method to take serial measurements on these 20 patients. Also, it is possible that some of the remaining 53 patients would be able to reach the top of the stool or beyond as their back condition changed. Having patients stand on a stool did not appear to cause any inconvenience or discomfort and none of the 73 patients reported any fear of falling off the stool

when forward bending.

This study is a first step in determining the usefulness of MFTF measurements. The FTF method has been referred to as a method for measuring lumbar flexion. Patients bending forward in an attempt to touch the floor can flex the hips and entire spine. They can also protract their scapulae, flex their shoulders, and extend their elbows, metacarpo-phalangeal and interphalangeal joints. Some authors argue that the FTF method should not be used to measure lumbar flexion because the FTF distance is dependent on motion occurring at a number of joints (Moll & Wright, 1987; Rae et al., 1984).

Individual differences in hip, spinal and upper extremity range of motion make the use of single FTF or MFTF measurements as a measurement of lumbar flexion questionable. However, some clinicians apparently assume that a decrease in the FTF distance on a patient with low back pain is primarily due to an increase in the patient's lumbar flexion (Ponte et al., 1984; Kendall & Jenkins, 1968).

Patients with low back pain are often able to reach farther towards the floor as they improve. The MFTF method, as described in this study, seems to be a reliable method for measuring an increase in a

patient's ability to reach toward the floor. If a patient with low back pain gains lumbar flexion there should be some change in the FTF distance. The ability of a patient to reach toward the floor is partially dependent on motion occurring in the lumbar spine as well as the hips and upper extremities. However, the relationship between changes in the FTF distance and changes in lumbar flexion has not been determined. The validity of using changes in MFTF measurements should be studied to determine if inferences on changes in lumbar flexion based on changes in the FTF distance are appropriate.

Suggestions for Further Study

To further determine the usefulness of MFTF measurements the issue of validity needs to be addressed. Further study needs to be performed to determine the degree of correlation between changes in the MFTF distance and changes in lumbar flexion.

Conclusions

The MFTF method, as used in this study, appears to be a very reliable method for measuring forward bending of patients with low back pain. MFTF measurements are easily and quickly obtained and the reliability of MFTF

measurements is apparently unaffected by an increase in pain during forward bending. However, the results of this study do not mean that the MFTF method can be used to assess changes in a patient's lumbar flexion. Further study needs to be performed to determine the degree of correlation between changes in the MFTF distance and changes in lumbar flexion. Once the validity of MFTF measurements for predicting the amount of lumbar flexion has been determined decisions can be made concerning the clinical usefulness of MFTF measurements.

Summary

The purpose of this study was to examine the reliability of measurements of forward bending obtained with the MFTF method. The FTF method was not used because it does not allow for measurement of patients who are able to touch the floor with their fingertips or patients who can reach beyond the level of the floor.

Repeated measurements were taken by six therapists on 73 patients with low back pain. Measurements were made over a short period of time. ICC's were calculated to determine intra-tester and inter-tester reliability.

The ICC values were high for intra-tester and inter-tester reliability. An increase in pain during forward bending did not appear to have any affect on reliability.

References

American Psychological Association (1974). Standards for educational and psychological tests. Washington, D. C.

Bartko, J. J. (1976). On various intraclass correlation reliability coefficients. Psychological Bulletin, 83, 762-765.

Bartko, J. J. & Carpenter, W. T. (1976). On the methods and theory of reliability. Journal of Nervous and Mental Disease, 163, 307-317.

Biering-Sorenson, F. (1984). Physical measurements as risk indicators for low-back pain. Spine, 9, 106-119.

Broer, M. R. & Galles, N. R. G. (1958). Importance of relationships between various body measurements in performance of the toe-touch test. The Research Quarterly, 29, 253-263.

Burdett, R. G., Brow, K. E. & Fall, M. P. (1986). Reliability and validity of four instruments for measuring lumbar spine and pelvic positions. Physical Therapy, 66, 677-684.

Burton, K. (1986). Regional lumbar sagittal mobility: Measurement by flexicurves. Clinical Biomechanics, 1, 20-26.

Buswell, J. (1982). Low back pain: A comparison of two treatment programmes. New Zealand Journal of Physiotherapy, 10, 13-17.

Buxton, D. (1957). Extension of the Kraus-Weber test. The Research Quarterly, 28, 210-217.

Calabro, J. J. (1982). Diagnosis of low back pain. In M. Stanton-Hicks 7 R. A. Boas (Ed.) Chronic low back pain. New York: Raven Press.

Cox, J. M. (1985). Low back pain: Mechanisms, diagnosis and treatment. (4th ed.) Baltimore: Williams & Wilkins.

Davies, J. E., Gibson, T. & Tester, L. (1979). The value of exercises in the treatment of low back pain. Rheumatology and Rehabilitation, 18, 243-247.

Doran, D. M. L. & Newell, D. J. (1975). Manipulation in treatment of low back pain: A multicentre study. British Medical Journal, 2, 161-164.

Dunham, W. F. (1949). Ankylosing spondylytis: Measurement of hip and spine movements. The British Journal of Physical Medicine, 12, 126-129.

Evans, D. P., Burke, M. S., Lloyd, K. N., Roberts, E. E. & Roberts, G. M. (1978). Lumbar spinal manipulation on trial part I - clinical assessment. Rheumatology and Rehabilitation, 17, 46-53.

Finneson, B. E. (1980). Low back pain. (2nd ed.). Philadelphia: J. B. Lippincott.

Frost, M., Stuckey, S., Smalley, L. A. & Dorman, G. (1982). Reliability of measuring trunk motion in centimeters. Physical Therapy, 62, 1431-1437.

Ghiselli, E. E., Campbell, J. P. & Zedeck, S. (1981). Measurement theory for the behavioral sciences. San Francisco: Freeman and Co.

Godfrey, C. M., Morgan, P. P. & Schatzker, J. (1984). A randomized trial of manipulation for low-back pain in a medical setting. Spine, 9, 301-305.

Haldeman, S., Gilles, J., Haldeman, J. & Patterson, C. (1975). Low back pain: A study of 50 patients on a group exercise program. Physiotherapy Canada, 27. 71-77.

Hoppenfeld, S. (1976). Physical examination of the spine and the extremities. New York: Appleton-Century-Crofts.

Kendall, P. H. & Jenkins, J. M. (1968). Exercises for backache: A double-blind controlled trial. Physiotherapy, 54, 154-157.

Kippers, V. & Parker, A. W. (1987). Toe-touch test: A measure of its validity. Physical Therapy, 67, 1680-1684.

Kraus, H. (1965). Backache, stress and tension: Cause, prevention and treatment. New York: Simon and Schuster.

Kraus, H. (1970). Clinical treatment of back and neck pain. New York: McGraw Hill Book Company.

Kubinec, J. (1977). Review article: low back dysfunction and its assessment. Physiotherapy Canada, 29, 73-80.

Loebl, W. Y. (1967). Measurement of spinal posture and range of spinal movement. Annals of Physical Medicine, 9, 103-110.

Lidstrom, A. & Zachrisson, M. (1970). Physical therapy on low back pain and sciatica: An attempt at evaluation. Scandinavian Journal of Rehabilitation Medicine, 2, 37-42.

Macrae, I. F. & Wright, V. (1969). Measurement of back movement. Annals of Rheumatic Disease, 28, 584-589.

Magee, D. J. (1987). Orthopedic physical assessment. Philadelphia: W. B. Saunders Company.

Maitland, G. D. (1986). Vertebral manipulation. (5th ed.). London: Butterworths.

Mayer, T. G., Tencer, A. F., Kristoferson, S. & Mooney, V. (1984). Use of noninvasive techniques for quantification of spinal range of motion in normal subjects and chronic low back dysfunction patients. Spine, 9, 589-595.

McKenzie, R. A. (1981). The lumbar spine: Mechanical diagnosis and therapy. New Zealand: Spinal Publications.

Mitchell, F. L., Moran, P. S., & Pruzzo, N. A. (1979). An evaluation and treatment manual of osteopathic muscle energy procedures. Valley Park, Mo.: Mitchell, Moran and Purzzo, Associates.

Moll, J. M. & Wright, V. (1971). Normal range of spinal mobility: An objective clinical study. Annals of Rheumatic Disease, 30, 381-386.

Moll, J. M. & Wright, V. (1973). The pattern of chest and spinal mobility in ankylosing spondylitis. Rheumatology and Rehabilitation, 12, 115-134.

Moll, J. M. H. & Wright, V. (1987). Measurement of spinal movement. In Jayson, M. (ed.), The lumbar spine and back pain. (3rd ed.). New York: Churchill Livingstone.

Moran, H. M., Hall, M. A., Barre, A. & Ansell, B. M. (1979). Spinal mobility in the adolescent. Rheumatology and Rehabilitation, 18, 181-185.

Ott, L. (1984). An introduction to statistical methods and data analysis (2nd ed.). Boston: Duxbury Press.

Paris, S. V. (1979). The spine: Etiology and treatment of dysfunction including joint manipulation, course notes. Stanley V. Paris.

Ponte, D. J., Jensen, G. J. & Kent, B. E. (1984). A preliminary report on the use of the Mckenzie protocol versus the Williams protocol. The Journal of Orthopedic and Sports Physical Therapy, 6, 130-139.

Portek, I., Pearcy, M. J., Reader. G. P. & Mowat, A. G. (1983). Correlation between radiographic and clinical measurement of lumbar spine movement. British Journal of Rheumatology, 32, 197-205.

Rae, P. S., Waddell, G. & Venner, R. M. (1984). A simple technique for measuring spinal flexion: Its use in orthopedic practice. Journal of the Royal College of Surgeons of Edinburgh, 29, 281-284.

Riddle, D. L., Rothstein, J. M. & Lamb, R. L. (1987). Goniometric reliability in a clinical setting: Shoulder measurements. Physical Therapy, 67, 668-673.

Rothstein, J. M., Miller, P. J. & Roettger, R. F. (1983). Goniometric reliability in a clinical setting: Elbow and knee measurements. Physical Therapy, 63, 1611-1615.

Rothstein, J. M. (1985). Measurement and clinical practice: Theory and application. In J. M. Rothstein (Ed.) Measurement in physical therapy. New York: Churchill Livingstone.

Schober, P. (1937). The lumbar vertebral column and backache. Muenchener Medizinesche Wochenschrift, 84, 336-338.

Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin, 86, 420-428.

Steel, R. G. D. & Torrie, J. H. (1960). Principles and procedures of statistics. New York: McGraw-Hill Book Company, Inc.

Taylor, J. & Twomey, L. (1980). Sagittal and horizontal plane movement of the human lumbar vertebral column in cadavers and in the living. Rheumatology and Rehabilitation, 19, 223-232.

Taylor, S. L. (1983). Low back pain assessment: part II - defining range of motion. Orthopedic Nursing, 2, 39-44.

Twomey, L. & Taylor, J. (1979). A description of two new instruments for measuring the ranges of sagittal and horizontal plane motions in the lumbar region. Australian Journal of Physiotherapy, 25, 201-203.

Winer, B. J. (1971). Statistical principles in experimental design (2nd ed.). New York: McGraw-Hill Book Company.

Williams, P. (1955). Examination and conservative treatment for disc lesions of the lower spine. Clinical Orthopedics, 5, 28-40.

APPENDIX A

SUBJECT CONSENT FORM

Appendix A

Subject Consent Form

The Reliability of Measurements Obtained with the Modified Fingertip-to-Floor Method on Patients with Low Back Pain

1. Introduction

The purpose of this study is to examine the reliability of the modified fingertip-to-floor method. Prior to being measured you will be asked a series of questions concerning your back problem. After answering the questions you will be instructed to stand on a 12 3/4 inch high stool and two physical therapists will individually assess your ability to flex your spine. A total of four measurements will be taken. Between measurements you will be allowed to move around the examination area for 15 to 30 second. This study is being performed by Captain Michael G. Gauvin, Department of Physical Therapy, Medical College of Virginia.

2. Benefits

There will be no direct benefit to you. Studies such as this need to be performed to assess the usefulness of evaluative tests used in the examination of patients with low back pain.

3. Alternative Therapy

NA

4. Risks, Inconveniences, Discomforts

The measurement session will require 5 to 10 minutes of your time. Although forward bending is commonly used in the examination of patients with low back pain, patients sometimes experience an increase in pain when performing this movement.

5. Cost of Participation

NA

_____

Patient Initials

6. Pregnancy

NA

7. Research Related Injury

"I understand that in the event of any physical and/or mental injury resulting from my participation in this research project, Virginia Commonwealth University will not offer compensation. I understand that my entitlements to medical and dental care and/or compensation in the event of injury are governed by federal laws and regulations, and if I desire further information I may contact the Malcolm Grow Medical Center Patient Affairs Office."

8. Confidentiality of Records

"Records of my participation in this study may only be disclosed in accordance with federal law, including the Federal Privacy Act, 5 USC 552a, and its implementing regulations. DD Form 2005 contains the Privacy Act Statement for records." All raw data will be reduced and analyzed as group data. If raw data is presented, only your subject number will identify you. Subject numbers are known only to the principal investigator.

9. Withdrawal

"The decision to participate in this study is completely voluntary on my part. No one has coerced or intimidated me into participating in this program. Captain Gauvin has adequately answered any questions I have concerning this study and Captain Gauvin will be available to answer questions during the study. I understand that I may withdraw from this study at any time without prejudice to my entitlements to care. I also understand that the investigator of this study may terminate my participation in this study if he feels this to be in my best interest."

_____ _____

Subject Signature Date

_____ _____

Witness Signature Date

APPENDIX B

DATA COLLECTION FORM

Appendix B

Data Collection Form

Subject Number _____ Age _____ Sex _____

Date _____

1. How long have you had the back pain you have currently?

2. Is this your first episode of pain? _____

   If not, when was your first episode?

3. What do you feel is the cause of your pain?

4. Is your pain constant or does it come and go?

5. Location (Shade in the appropriate areas)

6. What activities increase your pain?

7. What activities relieve your pain?

8. Do you have a history of:

   a. Rheumatoid arthritis

   b. Osteoarthritis of the back

   c. Ankylosing spondylitis

   d. Spinal tumor

   e. Spinal surgery

   f. Neurological disease

   g. Osteoarthritis of the hip

      Other hip problems

   h. Shoulder pain

   i. Other

9. Additional remarks

Referring Therapist #_____

Re-test Therapist #_____

Measurement #

1	_____cm	_____cm
2	_____cm	_____cm

Instructions to Patient

Consistent? Consistent?

Difference between therapists

Patient Positioning

Foot position

Consistent? Consistent?

	Meas. # 1	2	Meas. # 1	2
Knees straight	_____	_____	_____	_____
Knees flexed	_____	_____	_____	_____

Measurement Method

	Meas. # 1	2	Meas. # 1	2
Therapist stood				
In front of patient	____	____	____	____
Left of patient	____	____	____	____
Right of patient	____	____	____	____
Therapist				
Squatted	____	____	____	____
Forward bent	____	____	____	____
Therapist marked tape with				
Left thumb	____	____	____	____
Right thumb	____	____	____	____

During the measurement session the pain

Increased _____

Decreased _____

Did not change _____

APPENDIX C

INSTRUCTIONS TO THERAPIST

Appendix C

Instructions to Therapist

1. Instruct the patient to stand on the stool.

2. Instruct the patient to forward bend.

3. Use the blinded side of the tape measure to measure the distance between the tip of the right middle finger and the top of the stool. Mark the distance on the tape with your thumbnail. Do not look at the side of the tape measure with the centimeter scale.

4. Transfer the tape measure to the recorder. Keep the thumbnail at the point corresponding with the MFTF distance until the recorder takes the tape measure from you.

APPENDIX D

ARTICLE

THE RELIABILITY OF CLINICAL MEASUREMENTS OF FORWARD BENDING OBTAINED USING THE MODIFIED FINGERTIP-TO-FLOOR METHOD

Michael G. Gauvin

Dan L. Riddle

Jules M. Rothstein

This research was submitted while Captain Gauvin was a Physical Therapy Master's Candidate at the Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia. He is now chief, Physical Therapy, USAF Hospital, Tyndall Air Force Base, Florida.

Mr. Riddle is an Assistant Professor, Department of Physical Therapy, School of Allied Health Professions, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia.

Dr. Rothstein is an Associate Professor, Department of Physical Therapy, School of Allied Health Professions, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia.

Abstract

The purpose of this study was to examine the intra- and inter-tester reliability of measurements obtained using a modified version of the fingertip-to-floor (FTF) method of assessing forward bending. With the modified FTF (MFTF) method patients stand on a stool and forward bend so that measurements can be taken on patients who are able to reach beyond the floor. In this study randomly paired physical therapists took repeated MFTF measurements on 73 patients with low back pain. Intra-class correlation coefficients (ICC) were calculated for intra and inter-tester reliability. The ICC value for intra-tester reliability was 0.98 and the ICC value for inter-tester reliability was 0.95. The results of this study suggest that measurements of forward bending obtained on patients with low back pain using the MFTF method are highly reliable.

Physical therapists who examine and treat patients with low back pain usually assess the patient's ability to flex the lumbar spine. Many therapists probably assess a patient's lumbar flexion using observational methods and do not actually measure the amount of lumbar flexion. Some clinicians, however, feel that it is important to take measurements of lumbar flexion. Measurements of lumbar flexion are often used to help determine whether there may be a functional loss due to low back pain.[1] The selection of treatments for patients with low back pain may also be partially based on the assessment of lumbar flexion.[2] Measurements of lumbar flexion are also used to assess patient progress.[3]

A variety of methods for measuring lumbar flexion have been described. Angular measurements of lumbar flexion have been made using spondylometers[4], inclinometers[1,5] and standard goniometers.[6] Burton[7] used a draftman's flexible ruler to measure lumbar flexion. A tape measure has been used to take measurements over the lumbar spine and these measurements have been used to represent the amount of lumbar flexion.[8,9] All of these methods are similar in that the therapist palpates and marks spinous processes

and places the measurement instrument directly over the lumbosacral area.

Some clinicians have used the fingertip-to-floor (FTF) method for assessing lumbar flexion.[10] The FTF method differs from the methods previously discussed because FTF measurements are not taken directly over the lumbosacral area. FTF measurements are obtained by having the patient forward bend in an attempt to touch the floor with his fingertips. The distance between the tip of the middle finger and the floor is usually estimated, or measured with a tape measure[11] or yardstick.[12]

Some clinicians have taken serial FTF measurements on patients to assess the efficacy of treatments used for low back pain.[13,14,15,16] Therapists who use the FTF method to measure changes in lumbar flexion of patients with low back pain are essentially making assumptions. They are assuming that any change in the FTF distance is a result of a change in lumbar flexion because they are assuming that motion in the hips, thoracic spine and upper extremities remains unchanged between measurements.

One obvious problem with the FTF method is that measurements cannot be taken on patients who are able to touch the floor with their fingertips. Some authors

have modified the FTF method by having the subjects stand on a stool so that measurements could be taken when subjects were able to touch the floor with the fingertips.[11,12,17] Measurements are taken with the modified FTF (MFTF) method by having the subject forward bend in an attempt to touch the floor with the fingertips and then measuring the distance between the tip of the middle finger and the top of the stool. The MFTF method allows the therapist to take measurements on patients who are able to touch the floor or reach beyond the floor.

Physical therapists who use measurement methods, such as the MFTF method, to monitor patient progress attribute changes in the measurement to changes in the variable being measure. However, changes in measurements can also be partially due to measurement error. The amount of change due to error is dependent on the reliability of measurement.

The reliability of MFTF measurements taken on patients with low back pain has not been studied. Frost et al.,[11] Broer and Galles[12] and Buxton[17] concluded that measurements made with the MFTF method were reliable on healthy subjects. Broer and Galles[12] and Buxton[17] were interested in using the MFTF method on healthy subjects and chose a sample from the

population they wanted to study.

Frost et al. studied the reliability of measurements made with the MFTF method because they were interested in using the MFTF method to measure patients.[11] However, in selecting their sample they excluded subjects with a history of back surgery and subjects who had visited a physician because of low back pain within the five year period prior to their study. There are characteristics unique to a population that can affect reliability. Therefore, a sample from the population of interest should be selected when studying the reliability of measurements. Patients with low back pain might, for example, experience increased pain when asked to forward bend and touch the floor with their fingertips. Any change in a patient's pain with repeated tests might result in inconsistent measurements. Normal subjects typically have no pain during forward bending. Therefore, the reliability of measurements made with the MFTF method on patients with low back pain might be different than the reliability of measurements taken on normal subjects.

The purpose of this study was to assess the intra-tester and inter-tester reliability for measurements of forward bending taken with the MFTF method on patients

with low back pain. The questions in this study were:

1. What is the intra-tester reliability of measurements obtained with the MFTF method on patients with low back pain?
2. What is the inter-tester reliability of measurements obtained with the MFTF method on patients with low back pain?

METHOD

Subjects

Subjects for this study were 73 patients referred for treatment of low back pain to the Department of Physical Therapy, Malcolm Grow Medical Center, Andrews Air Force Base, Maryland. Patients were included in the study if assessment of the amount of lumbar flexion was considered by the therapist to be an appropriate part of the patient's examination. All patients included in this study were at least 18 years of age. Patients were asked to participate if the patient's therapist felt the patient could tolerate repeated forward bending. The subjects in this study were 47 males and 26 females ranging in age from 18 to 73 years. The mean age was 43.5 years with a standard deviation of 13.9 years. All patients were asked to read and sign a consent form prior to inclusion in the study.

Testers

Testers for this study were six physical therapists on the staff of the Department of Physical Therapy, Malcolm Grow Medical Center, Andrews Air Force Base, Maryland. The therapists' experience ranged from 6 months to 30 years. All testers were given a description of the MFTF method prior to the beginning of the study by one of the investigators (MG). The description was limited to instruction in the use of the tape measure and the stool that were used in the study. Testers were not given specific instructions on how to position patients or how to instruct patients to bend forward. None of the testers reported they routinely took FTF or MFTF measurements.

Prior to the beginning of the study each therapist was assigned a number. These numbers and a table of random numbers were used to make a random list for each therapist. These lists were used to randomly choose a second therapist for each patient measured.

Instrumentation

A cloth tape measure* marked in .10 cm increments was used to obtain measurements with the MFTF method.

---

*Cloth Tape Measure, Dritz Corporation, Spartanburg, SC, 29304

The reverse side of the tape measure was covered with silver duct tape to prevent the therapists from seeing the numbers on the tape measure and to prevent examiner bias. To assure that the length of the tape measure did not change with repeated use one of the investigators (MG) checked the accuracy of the tape measure twice a day by comparing the cm scale on the tape to a metal meter stick. The length of the tape measure did not change during the study.

Procedure

An investigator (MG) was notified whenever a therapist (called the referring therapist) identified an appropriate patient for the study. The investigator described the study to the patient and asked the patient to read and sign the consent form. After the consent form was signed the referring therapist used his random number list to identify another therapist (called the re-test therapist) who would take a second set of measurements on the patient.

The referring therapist instructed the patient to stand on a 12 3/4 inch (32.4 cm) high stool and bend forward. Using the blank side of the tape measure the therapist measured the distance between the tip of the patient's right middle finger and the top of the stool. The therapist marked the distance on the tape measure

with his thumbnail (Figure 1). The therapist then held the tape measure so that the investigator (MG) could see the marked side of the tape measure. The investigator used his thumbnail to mark the point on the cm scale that corresponded with the referring therapist's thumbnail and then took the tape from the referring therapist (Figure 2). The referring therapist was not allowed to see the marked side of the tape measure and was not informed of the measurement.

The MFTF distance was recorded to the nearest .10 cm. A positive value was recorded if the patient was unable to reach the top of the stool. A negative value was recorded if the patient was able to reach beyond the top of the stool. A zero was recorded if the patient reached the top of the stool. After the first measurement was taken the patient was allowed to step off the stool and move around the measurement area for 15 to 30 seconds. The patient then stepped back on the stool and the referring therapist took a second measurement. After the referring therapist took two measurements the re-test therapist was notified and took two measurements using the same method as the referring therapist. The re-test therapist was not informed of the measurements obtained by the referring therapist.

Figure 1. Therapist taking MFTF measurement.

Figure 2. **Therapist transferring the tape measure to the recorder.**

Data Analysis

Intraclass correlation coefficients (ICC [1,1][18]) were calculated to assess intra-tester and inter-tester reliability. ICC(1,1) was chosen because it is the appropriate form of ICC to use when assessing the reliability of single measurements taken by random pairs of testers. ICC(1,1) is also the most conservative form of ICC.

The ICC for intra-tester reliability was calculated by comparing the first and second measurements taken by all therapists. Therefore, there were 146 paired measurements used to determine the intra-tester reliability of MFTF measurements. The ICC for inter-tester reliability was calculated by comparing the first measurement taken by the referring therapist with the first measurement taken by the re-test therapist. There were 73 paired measurements used to determine the inter-tester reliability of MFTF measurements.

RESULTS

The ICC for intra-tester reliability of all paired measurements was .98. The ICC for inter-tester reliability was .95 (Table 1).

DISCUSSION

The results of this study were similar to those of

Table 1

Intraclass Correlation Coefficients for Reliability

	N	ICC
Intra-tester Reliability	146	.98[a]
Inter-tester Reliability	73	.95[b]

[a] This ICC was calculated by comparing the first and second measurements of the referring and re-test therapist.

[b] This ICC was calculated by comparing the first measurement taken by the referring therapist with the first measurement taken by the re-test therapist.

Frost et al.,[11] who reported high intra-tester and inter-tester reliability for MFTF measurements taken on healthy subjects. However, there are a number of methodological differences between the two studies.

Frost et al. studied only healthy subjects. Therefore, the reliability values reported by Frost et al. are of limited value to clinicians because the reliability of MFTF measurements obtained on healthy subjects might differ from the reliability of MFTF measurements obtained on patients. In the present study 73 patients with low back pain were measured because the population of interest was patients with low back pain.

Selecting a sample from the population of interest is important because each population has characteristics that might affect the reliability of measurements. Patients with low back pain commonly have an increase in pain when forward bending. Frost et al. were unable to examine the effect of pain on the reliability of MFTF measurements because they examined normal subjects.

In the present study 22 of the 73 patients stated that they had an increase in pain during forward bending. To determine if an increase in pain had any effect on the reliability of measurements a posteriori

analyses were performed. Separate ICC's were calculated for repeated measurements taken on patients who had an increase in pain with forward bending and patients who had no change in pain with forward bending. The ICC for patients who had an increase in pain was .97 and the ICC for patients who had no change in pain was .94. The ICC's for inter-tester reliability were essentially the same for patients who had no change in pain during forward bending and patients who had an increase in pain.

An increase in pain during forward bending apparently did not influence reliability but it may have affected the patient's ability to forward bend. The results of an a posteriori t-test revealed that the mean of the first MFTF measurements (20.65 $\pm$ 14.5 cm) for patients who had an increase in pain when forward bending was significantly greater ($p<.05$) than the mean of the first MFTF measurements (12.6 $\pm$ 12.9 cm) for patients who had no increase in pain when forward bending). The results of the t-test suggest that patients who report an increase in pain when forward bending may have a decrease in the ability to forward bend when compared to patients who do not report an increase in pain during the test.

Another difference between the present study and the study by Frost et al.[11] was the instructions the therapists gave to the subjects. In the study by Frost et al. the therapists who took the measurements used a specific set of instructions. In the present study the therapists were allowed to use their own instructions. No pair of therapists gave the same instructions to the patients and the instructions varied greatly. One therapist gave very brief instructions, such as "bend forward as far as you can." Another therapist gave very specific instructions, telling the patient to "bend forward by rolling the spine, bending at the hips and keeping the knees straight."

The therapists in the present study allowed the patient to stand in the position that was most comfortable for the patient. Most patients stood with their feet approximately shoulder width apart. The therapists were allowed to give their own instructions to the patients and patient positioning was not standardized in order to more closely replicate clinical procedure. The results of this study reveal that standardized instructions and patient positioning are not required for MFTF measurements to be highly reliable.

The results of this study indicate that the intra-tester and inter-tester reliability of MFTF measurements taken on patients with low back pain is very high. Data collected in this study also indicate that the MFTF method may be preferable to the FTF method for therapists who take serial measurements on patients. Out of the 73 patients measured in this study 20 (27%) were able to reach the top of the stool or beyond the top of the stool. A therapist would need to use the MFTF method to take measurements on these 20 patients. Also, it is possible that some of the remaining 53 patients would be able to reach to top of the stool or beyond as their low back pain changed. Having patients stand on a stool did not appear to cause any inconvenience or discomfort. None of the 73 patients reported any fear of falling off the stool when bending forward.

This study has shown that MFTF measurements taken on patients with low back pain can be highly reliable. Therefore, this study is a first step in determining the usefulness of MFTF measurements. However, the usefulness of MFTF measurements depends on the validity of measurements obtained with the MFTF method.

The FTF method has been referred to as a method for measuring lumbar flexion, but the FTF distance is

not just a measurement of flexion of the lumbar spine. When a patient bends forward in an attempt to touch the floor with his fingertips, he can flex the hips and entire spine. He can also protract the scapulae, flex the shoulders, and extend the elbows, metacarpophalangeal and interphalangeal joints. Some authors argue that the FTF method should not be used to measure lumbar flexion because the FTF distance is dependent on motion occurring at a number of joints.[19,20]

Biering-Sorenson[21] compared single FTF measurements with measurements of lumbar flexion obtained with the skin distraction method.[8] Measurements were taken on 479 female subjects and 449 male subjects. Biering-Sorenson reported a correlation coefficient of 0.22 for females and -0.35 for males. Biering-Sorenson's results reveal that there may not be a strong linear relationship between the FTF distance and the amount of lumbar flexion. However, because Biering-Sorenson used the FTF method and not the MFTF method he was not able to measure subjects who were able to reach the floor or beyond. He merely assigned these subjects a value of zero. If Biering-Sorenson would have used the MFTF method the results may have been different.

In our opinion, individual differences in hip, spinal and upper extremity range of motion make the use of a single FTF measurement as a measurement of lumbar flexion questionable. However, most therapists who assess forward bending of patients with low back pain probably do not use a single FTF measurement as a measurement of lumbar flexion but take serial measurements over time. Any change in the FTF distance is then attributed to a change in lumbar flexion. Patients with low back pain are often able to reach farther towards the floor as they improve. If a patient with low back pain gains lumbar flexion there should be some change in the FTF distance because the ability of a patient to reach toward the floor is partially dependent on motion occurring in the lumbar spine.

The degree of correlation between changes in the FTF distance and changes in lumbar flexion has not been determined. Validity studies need to determine if the FTF distance can be used to reflect changes in lumbar flexion.

One method for studying the validity of changes in the FTF distance would be to compare serial FTF measurements with serial measurements of lumbar flexion obtained from radiographs. Because some patients can

reach beyond the level of the floor the MFTF method would be preferable to the FTF method in a validity study. The MFTF method as described in this study seems to be reliable and therefore appropriate for any future study addressing the validity of MFTF measurements.

CONCLUSIONS

The MFTF method, as used in this study, appears to be a very reliable method for measuring forward bending of patients with low back pain. MFTF measurements are easily and quickly obtained and the reliability of MFTF measurements is apparently unaffected by an increase in pain during forward bending. However, the results of this study do not provide evidence as to whether the MFTF method can be used to assess changes in a patient's ability to flex the lumbar spine. Further study needs to be performed to determine the degree of correlation between changes in the MFTF distance and changes in lumbar flexion. Once the validity of MFTF measurements for predicting the amount of lumbar flexion has been determined, decisions can be made concerning the clinical usefulness of MFTF measurements.

References

1. Mayer TG, Tencer AF, Kristoferson S, Mooney V: Use of noninvasive techniques for quantification of spinal range of motion in normal subjects and chronic low back dysfunction patients. Spine 9:589-595, 1984
2. McKenzie RA: The lumbar spine: mechanical diagnosis and treatment. New Zealand, Spinal Publications, 1981
3. Maitland GD: Vertebral manipulation. (5th ed). London, Butterworth, 1986
4. Twomey L, Taylor: A description of two new instruments for measuring the ranges of sagittal and horizontal plane motion in the lumbar region. Aust J Physiother 25:201-203, 1979
5. Loebl WY: Measurements of spinal posture and range of spinal movement. Ann Phys Med 9:103-110, 1967
6. Cox JM: Low back pain: Mechanism, diagnosis and treatment. (4th ed). Baltimore, Williams & Wilkins, 1985
7. Burton K: Regional lumbar sagittal mobility. Clin Biomech 1:20-26, 1986

8. Macrae IF, Wright V: Measurement of back movement. Ann Rheu Dis 28:5844-589, 1969

9. Schober P: The lumbar vertebral column and backache. Muenchener Medizinesch Worchenschrift 84:336-338, 1937

10. Hoppenfeld S: Physical examination of the spine and the extremities. New York, Appelton-Century-Crofts, 1976

11. Frost M, Stuckey S, Smalley LA, Dorman G: Reliability of measuring trunk motion in centimeters. Phys Ther 62:1431-1437, 1982

12. Broer MR, Galles NRG: Importance of relationships between various body measurements in performance of the toe-touch test. Res Q 29:253-263, 1958

13. Ponte DJ, Jensen GH, Kent BE: A preliminary report on the use of the McKenzie protocol versus the Williams protocol. JOSPT 6:130-139, 1984

14. Kendall PH, Jenkins JM: Exercises for backache: A double-blinded controlled trial. Physiother 54:154-157, 1968

15. Lidstrom A, Zachrisson M: Physical therapy on low back pain and sciatica: An attempt at evaluation. Scand J. Rehab Med 2:37-42, 1970

16. Haldeman S, Gi˙les J, Haldeman J, Patterson C: Low back pain: A study of 50 patients on a group

exercise program. Physiother Can 27:71-77, 1975

17. Buxton D: Extension of the Kraus-Weber test. Res Q 28:210-217, 1957

18. Shrout PE, Fleiss JL: Intraclass correlations: Uses in assessing rater reliability. Psych Bull 86:420-428, 1979

19. Rae PS, Waddell G, Venner RM: A simple technique for measuring spinal flexion: Its use in orthopedic practice. J Roy Coll Surg Edingburgh 29:281-284, 1984

20. Moll JMH, Wright V: Measurement of spinal movement. In Jayson M (ed). The lumbar spine and back pain (3rd ed). New York, Churchill Livingstone, 1987

21. Biering-Sorenson F: Physical measurements as risk indicators for low back pain. Spine 9:106-119, 1984

ACKNOWLEDGEMENT

The authors would like to thank the therapists at Malcolm Grow Medical Center for their assistance in data collection.

Vita

Michael G. Gauvin, Captain, USAF, BSC [redacted] He received a Bachelor of Science degree in Physical Therapy from the University of Minnesota in 1981 and a Master of Science degree in Health Science from Chapman College in 1985. He has practiced as a physical therapist in the United States Air Force for eight years and is a member of the American Physical Therapy Association.