2012

Development of an ankle function model

Kelli R. Snyder
University of Northern Iowa

DEVELOPMENT OF AN ANKLE FUNCTION MODEL

An Abstract of a Dissertation

Submitted

in Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

Approved:

____
Dr. Todd Evans, Committee Co-Chair

____
Dr. Sam Lankford, Committee Co-Chair

____
Dr. Michael J. Licari
Dean of the Graduate College

Kelli R. Snyder

University of Northern Iowa

July, 2012

ABSTRACT

Lateral ankle sprains (LAS) represent the most common orthopedic injury treated by athletic trainers, yet there is no consensus regarding the measurement of the effectiveness of care (e.g. outcomes). Further compounding this problem is the repetitive nature of LAS, often referred to as "functional" or "chronic" ankle instability (FAI, CAI, respectively). Although a common entity in practice and research, FAI is inconsistently defined and assessed. Essentially, athletic trainers are neglecting to address one of the most important issues identified in their profession (outcomes) for the most common injury for which they provide care (LAS). Therefore, the purpose of this research was to develop a model for ankle function which allows researchers and clinicians to better assess patient outcomes following ankle injury.

The Delphi method involving one round of telephone interviews and four rounds of online surveys with 16 content experts was used to gather feedback regarding the definition, functional characteristics, and assessment techniques of a healthy ankle, an unhealthy ankle, and FAI/CAI. Additionally, 10 female and nine male elite athletes were interviewed to determine their functional limitations following foot and ankle injury and their perceptions of an ankle outcomes instrument.

A consensus was reached for the definition of a healthy ankle: "A person with a healthy ankle(s), either through self-report or clinician measurement, presents with full functional capacity and participation status, without pathology, pain, or impairments relative to the ankle(s)." Additionally, 11 functional characteristics of a healthy ankle and 32 functional characteristics of an unhealthy ankle, along with 33 assessment

techniques were agreed upon. The definition of FAI was established as: "a recurrent sense of giving way of the ankle." However, a consensus was not reached for the definition of chronic ankle instability. Athletes consistently indicated that the items on the instrument need to be difficult and sport-specific.

By creating a model of ankle health, and identifying the functional limitations that elite level athletes suffer due to foot and ankle injury, this study provides the framework for establishing accurate outcomes assessment of individuals with high levels of physical ability, and thereby facilitates the implementation of evidence-based practice in athletic training.

DEVELOPMENT OF AN ANKLE FUNCTION MODEL

A Dissertation

Submitted

in Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

Approved:

Dr. Todd Evans, Co-Chair

Dr. Sam Lankford, Co-Chair

Dr. Peter Neibert, Committee Member

Dr. Windee Weiss, Committee Member

Dr. Gayle Rhineberger-Dunn, Committee Member

Kelli R. Snyder

University of Northern Iowa

July, 2012

DEDICATION

I would like to dedicate this dissertation to my sons, Joey and Colin, and to my husband, Andy. Their love and support have made this dream a reality. I could never have accomplished this goal without my husband's unconditional love and support.

ACKNOWLEDGEMENTS

I would like to thank Dr. Todd Evans for his unwavering guidance and mentorship throughout this process, and for his endless enthusiasm and encouragement. I truly cannot thank him enough.

I would like to thank Dr. Sam Lankford for his willingness to be my co-chair and for his constant guidance. He was always very quick to answer my questions.

I would like to thank Dr. Peter Neibert for providing his expertise and positive attitude. He was always willing to answer my questions and attend spur of the moment meetings.

I would like to thank Dr. Windee Weiss for her attention to detail and encouragement throughout this process. On many days, she was there to give me the push I needed to keep going.

I would like to thank Dr. Gayle Rhineberger-Dunn for her careful reading, thoughtful insights, and for prompting me think about this project from another perspective.

I would like to thank Tricia Haak for so generously offering her time and assistance to me and this project.

I thank my extended family and friends for all of their love and support throughout this process. Without their willingness to help, or cheering me on the whole way, I could not have accomplished this goal.

Finally, I would like to thank the Provost's Office for awarding me a Pre-Tenure Research Fellowship to assist with this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Evidence-based practice (EBP) is one of the most important and prominent issues in the athletic training profession, as identified by the National Athletic Trainers' Association Research and Education Foundation (NATAREF, 2011; Steves & Hootman, 2004). EBP is an approach in which health care decisions are based upon the best available evidence, the practitioner's clinical expertise, and the patient's values and expectations (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996). In order to establish EBP, however, patient outcomes must first be routinely documented and assessed during the daily health care practice. Outcomes assessment is the process of monitoring patients' changes in health status that result from treatments provided by athletic trainers or other health care practitioners (Donabedian, 1980). The NATAREF has clearly identified outcomes assessment as a research priority (NATAREF, 2011). Although there are various factors limiting outcomes assessment in athletic training, several consistent barriers exist. Of paramount concern is identifying outcomes that are oriented to the unique clientele served by athletic trainers, as well as selecting the appropriate outcome instruments through which they should be assessed.

There is not a more vivid example of the barriers to consistently assessing outcomes than in the treatment of ankle injuries. Specifically, lateral ankle sprains (LAS) represent the single most common athletic injury (Hootman, Dick, & Agel, 2007), and have the highest re-injury rate among all injuries (Borowski, Yard, Fields, & Comstock, 2008; Hootman et al., 2007; Sankey, Brooks, Kemp, & Haddad, 2008), yet outcomes that substantiate their treatment are limited. One barrier is the availability of valid outcome instruments; of the more than 40 instruments used for the ankle and foot (Button & Pinney, 2004), none have been validated for athletic training clientele. This potentially creates an issue when these instruments are used to assess patients with high levels of function, as seen in athletes. The athletes appear healthy, even though they may still have functional limitations which may not be detected if the items on the instruments are too easy; that is, the instruments fail to address the higher ability levels of physically active individuals. One example of an instrument with merit, but lacking the difficulty necessary for individuals with high levels of physical ability, is the Foot and Ankle Disability Index (FADI; Martin, Burdett, & Irrgang, 1999).

Originally developed in 1999 to assess restrictions in function that result from foot and ankle injury, the FADI was intended for use for patients with a variety of ability levels. To develop the initial items, a literature review was conducted, patient feedback was gathered, other outcomes instruments were reviewed, and expert opinions were gathered from the American Physical Therapy Associations's Foot and Ankle Special Interest Group (Martin et al., 1999). The FADI-Total was then divided into two subscales; activities of daily living (ADL; 22 activity items, 4 pain items) and sports-related activities (Sport; 8 items). The FADI-Total, therefore, consists of 34 items. Although this instrument has merit, it has not been validated in a group of high ability individuals, and has been shown to have a ceiling effect (i.e., the respondents achieve a maximum score, although they are not fully recovered, and it is no longer possible to monitor progress; Schlitz, Evans, & Ragan, 2009). A logical solution to this problem is to create more difficult items that could be added to an instrument, such as the FADI, to make it appropriate for use with individuals who possess high physical ability. However, any changes to the FADI items should include consultation with experts in the area of ankle function, treatment, and general outcomes assessment and research, as well as injured athletes.

Further compounding the problem of assessing outcomes following LAS is the repetitive nature, which has led to the identification of a phenomenon called "chronic" or "functional" ankle instability (FAI, CAI, respectively; Freeman, Dean, & Hanham, 1965; Hertel, 2002; Tropp, 2002). Although a common entity in practice and research, FAI is inconsistently defined and assessed (Howe-Weichers, 2007). The first inconsistency regarding FAI is its definition. Despite widespread assimilation, and overwhelming support for its' existence, there is currently no consensus regarding FAI's definition (Howe-Weichers, 2007). The second is that there is no consensus regarding the identification or characteristics present for those who have this phenomenon. Researchers have classified the existence of FAI inconsistently and clinicians have diagnosed it arbitrarily. To substantiate the existence of any injury or condition, it must be definable and have a measurable impact on function. Neither of these is true regarding FAI. This has left discrepancies in the research because no two researchers classify the condition using the same criteria (Howe-Weichers, 2007). Therefore, cross-study comparisons are impossible and the true impact of the condition is left largely unknown. Without consistent classification, outcomes cannot be accurately measured. The lack of standard outcomes and the inconsistencies in identification of repetitive LAS has further compounded the reporting of ankle outcomes. Therefore, athletic trainers are neglecting to address one of the most important issues identified in their profession (outcomes) for the most common injury for which they provide care (LAS).

In following the model of the World Health Organization (WHO), the first logical step towards outcomes assessment following LAS would be to develop a model for a healthy ankle, establish a clear definition for the recurring ankle sprain phenomenon or FAI, followed by the identification of the functional outcomes that should be assessed following LAS (World Health Organization, 2002). This can be determined by interviews with content experts and individuals with recent history of an ankle sprain. Once ankle "health" and the appropriate outcomes that address high levels of physical ability are identified, new and more difficult items can be added to an established ankle outcome instrument. Then the modified instrument can be validated for athletic training clientele (Evans & Lam, 2011).

Statement of the Purpose

The collective purpose of this research is to develop a model for ankle function which allows researchers and clinicians to better assess patient outcomes for ankle injury in individuals with high levels of physical ability.

Significance of the Study

This study builds a foundation for future research that will assist the development of new, more difficult items which could be added to an existing ankle outcomes instrument, as identified through interviews with content experts and elite athletes. Developing items the potentially represent greater difficulty will allow clinicians and researchers to more accurately assess ankle outcomes following injury in patients with high levels of physical ability.

CHAPTER II

REVIEW OF LITERATURE

The commonly accepted approach for providing optimal health care is through the implementation of evidence-based-practice (EBP). EBP is an approach in which health care decisions are based upon the best available evidence, the practitioner's clinical expertise, the patient's values and expectations, and the routine assessment of the patient's health status (Sackett et al., 1996). The component of EBP that involves monitoring the health status of patients throughout the delivery of health care is called outcomes assessment (Jette, 1995). Unfortunately, the profession of athletic training is behind in regards to implementing EBP and ensuring that services provided to patients are supported by all of the components of evidence. More specifically, athletic trainers do not routinely track patient outcomes in their daily patient care. Therefore, an important step in establishing EBP within athletic training is the routine assessment and documentation of patient outcomes during the daily practice.

In general, outcomes assessment is the process of monitoring patients' changes in health status that result from treatments provided by athletic trainers or other health care practitioners (Donabedian, 1980). To accomplish this however, valid and efficient outcome instruments must be available to measure the health status of athletic training clientele. One of the important barriers limiting outcomes assessment in athletic training is the availability of outcome instruments that are appropriate for athletic training clientele which have a high level of physical ability.

There is not a more vivid example of the barriers to consistently assessing outcomes than in the treatment of ankle injuries. Specifically, lateral ankle sprains (LAS) represent the single most common athletic injury (Hootman et al., 2007), and have the highest re-injury rate among all injuries (Hootman et al., 2007; Sankey et al., 2008), yet outcomes that substantiate their treatment are limited. Therefore, athletic trainers are neglecting to address one of the most important issues identified in their profession (outcomes) for the most common injury for which they provide care (LAS).

The purpose of this literature review is to delineate the necessity of outcomes assessment and review the current issues with outcome instruments. Specifically, it will provide an overview of outcomes and explain how an instrument should be properly created and selected for clinical use. It will provide a description of ankle function and the conundrum of assessing it properly, and provide a framework for addressing the problems pertaining to ankle outcome instruments appropriate for athletic training clientele. Lastly, it will explain how these issues will be addressed.

Measuring Health Care Outcomes

Athletic trainers must closely monitor the course of injury or disease from onset to return-to-activity to ensure that the highest quality health care interventions are provided and their effectiveness is being monitored. The accepted approach for delivering the highest level of patient care is EBP. By definition, EBP is "the integration of best research evidence with clinical expertise and patient values" (Sackett et al., 1996). Incorporation of evidence into clinical practice allows clinicians to provide the best possible care to patients (Steves & Hootman, 2004). By urging clinicians to seek the most current literature, think critically, and have an open mind to change, evidence-based practice encourages the use of the most scientifically supported treatments (Steves & Hootman, 2004).

The evolving recognition of the importance of EBP has led to the increased implementation of outcomes assessment as a means of monitoring the effects of treatment interventions (Jette, 1995). Outcomes assessment is the process of monitoring changes in patients' health status during health care interventions delivered by athletic trainers or other health care providers (Donabedian, 1980) and is an integral component of the evidence-based triad (Evans & Lam, 2011). Donabedian (1988) established a framework for the establishment of evidence-based practice, which includes structure, process, and outcomes evaluation. Structure evaluation includes assessment of the resources necessary for proper patient care. Process evaluation includes examination of the quality of provided interventions. Lastly, outcomes evaluation involves the assessment of the end result of health care (Donabedian, 1988), which allows clinicians to monitor and evaluate the quality of their interventions (Evans & Lam, 2011).

Although evidence-based practice (EBP) is central to decision-making in most allied health care professions, the profession of athletic training has been slow, even resistant, to incorporate EBP into clinical practice and has yet to establish it as a routine component of their clinical practice (Evans & Lam, 2011), even though the National Athletic Trainers' Association Research and Education Foundation has clearly identified outcomes assessment as a research priority (NATAREF, 2011). Clinicians fear that EBP will hinder decision-making freedom and relies solely on research evidence to dictate patient care, but as Sackett et al. (1996) point out, EBP is not "cookbook" practice. It involves clinician experience, research evidence, and patient values; none of which is sufficient without the others (Sackett et al., 1996). Athletic training has been slow to embrace EBP also because, for the most part, their services are not reimbursed by third party payers and therefore, have not been required to show accountability for their treatments rendered. However, athletic trainers are currently fighting for third party reimbursement, making it vital for their practice to be supported by evidence. Additionally, as a profession, it is imperative to be able to support the practice of athletic training with evidence in order to gain respect from other health care professionals, and to encourage patients to seek medical care from athletic trainers (Steves & Hootman, 2004).

One of the major barriers in the transition to EBP is the selection of appropriate outcomes instruments (Evans & Lam, 2011; Steves & Hootman, 2004). There are countless outcomes instruments available from which health care providers must choose, which complicates the decision-making process. It takes time to sift through each instrument to determine which is the most appropriate for each case, and which are supported by validity measures. A vivid example of the barriers to selecting outcomes instruments is the treatment of ankle injuries. Specifically, lateral ankle sprains (LAS) represent the single most common athletic injury (Hootman et al., 2007), and have the highest re-injury rate among all injuries (Borowski et al., 2008), yet outcomes that substantiate their treatment are limited. A key limitation in the establishment of appropriate ankle outcomes is the availability of valid outcome instruments; of the more than 40 instruments used for the ankle and foot (Button & Pinney, 2004) none have been validated for high ability patients. This potentially creates an issue when these instruments are used to assess patients with high levels of function. The patients appear healthy, even though they may still have functional limitations, which may not be detected if the items on the instruments are too easy. In other words, the items on the instruments fail to address the higher ability levels of physically active individuals.

Defining Health Status

One of the most important, and often the most challenging, steps in the process of assessing health care outcomes is identifying the element of health that is of concern and finding an instrument that can adequately quantify it. These elements, which are often referred to as components or dimensions of health, will then be monitored to determine the effectiveness of the health care interventions, track recovery, and establish if interventions are still necessary, which ultimately addresses the end result of health care. This can be addressed on several different levels. The outcome itself can be the presence of a disease or injury. For example, the outcomes related to preventative techniques in athletic training are the presence, or absence, of the injury that is trying to be prevented. A more common, and somewhat challenging, element of health that is often addressed in health care outcomes is level of health. A patient's level of health has been described as function (Nagi, 1965), activity limitation (World Health Organization International Classification of Functioning, Disability, and Health; World Health Organization, 2002), disability (Nagi, 1965), impairment (Nagi, 1965; World Health Organization, 2002), and participation status (World Health Organization, 2002), among other descriptors. These levels of health status have been addressed, identified, and categorized in various frameworks.

Models have been established as frameworks to provide understanding of the disablement process and to provide a common language that can be shared by practitioners throughout the world of health care. In general, these models provide a "conceptual framework" (Snyder et al., 2008), whereas outcomes assessment provides the means, for evidence-based practice (Jette, 2006; Snyder et al., 2008). The primary importance of these models is that they focus more expansively on the disability to include assessment of the patient's social well-being, rather than focusing primarily on the physical limitations created by the disability (Jette, 2006; Snyder et al., 2008).

The original model was introduced in 1965 by Nagi. Nagi's disablement model included four components: active pathology, impairment, functional limitations, and disability. Active pathology describes the injury at the cellular level (i.e. anterior talofibular ligament tear). Impairment refers to dysfunction at the tissue, organ, or system level (i.e. ankle pain while ambulating, limited dorsiflexion range of motion). Functional limitations describe how the pathology affects the person as a whole, with particular concentration on social function and activities of daily living (i.e. inability of a basketball player to run full speed up and down the court). The fourth component, disability, refers to the incapacitation of a person to perform the desired levels of activity, either socially or personally (i.e. loss of starting position on the basketball team, leading to role-identity crisis and loss of self-worth). It is important to note that two individuals with the same injury may experience vastly different disabilities. For instance, one individual may continue their normal social activities, despite their injury; whereas another may entirely withdraw from social interaction. (Snyder et al., 2008)

Although Nagi's model set the groundwork for disablement models, and was used as the primary model for many years, other models have emerged and become widely used. Most recently, the World Health Organization's International Classification of Functioning Model (WHO ICF) has been developed and refined to provide a common international language to describe disability which spans all realms of health care, and spans time (World Health Organization, 2002). Although the structure of this model is quite complex, there are key terms and concepts that should be identified. Impairment refers to loss in function or structural abnormalities at the body organ or system level (i.e. dorsiflexion range of motion, tibialis anterior muscle strength; Jette, 1995). Participation, on the other hand, describes the multifaceted relationships between the pathology from which the person is suffering and the environment in which they function, including social functioning (i.e., inability to attend a business meeting conducted on a golf course due to an ankle sprain; World Health Organization, 2002). Unfortunately, athletic trainers commonly neglect to account for the disability of their patients and focus primarily on the impairments. Both, in fact, are essential to the proper diagnosis, treatment, progression, and return-to-activity decisions made by the clinician. It is imperative that athletic training embraces the use of disablement models as the conceptual framework for EBP in order to achieve sustainability as a health care profession (Parsons, Valovich McLeod, Snyder, & Sauers, 2008).

Quantifying Health Status

Not only is it critical to accurately identify health status, the level of health status must also be quantified. The measurement tools used to quantify changes in health status are called outcome instruments. One critical ingredient in tracking health status is the outcome instrument. Although selecting the appropriate instrument would seem to be a simple step in the outcome assessment process, it can ultimately be a frustrating and confusing quest; one that clinicians and researchers often find difficult because of an overwhelming number of choices and too much or too little supporting psychometric data. However, selecting an inadequate or inappropriate instrument can render even the best designed outcome systems and studies meaningless.

Types of Outcome Instruments

In the most general sense, health status, or outcome, instruments are classified as either clinician-reported or patient/self-reported (McLeod et al., 2008). Clinician-reported outcomes, often mistakenly referred to as objective measurements, are measurements taken by the clinician such as ligament laxity, joint range of motion, or muscle strength. These measurements focus on the evaluation of the injury or condition, as well as on impairments suffered due to injury (McLeod et al., 2008). Clinician-reported outcomes are important for tracking injury/disease progression, intervention effectiveness, and providing appropriate care (Denegar, Vela, & Evans, 2008). Although these assessments are a necessary part of the evaluation process, they do not include assessment of the patient's perception of the injury or condition. Therefore, it is imperative for clinicians to incorporate patient-reported outcomes into their injury evaluation, treatment, and return to activity decision-making, in addition to the traditional clinician-reported measures. Patient-reported outcomes instruments provide a means of gathering patients' perceptions of the impact of the injury or condition and measurement of the effectiveness of treatment intervention (McLeod et al., 2008; Valderas et al., 2008; Wiklund, 2004) and can be used both in the clinical and research settings. Patient-reported outcomes help to bridge the information gathered in the clinician-reported portion of the assessment to how the patient perceives the progress of the injury or condition; thereby adding important information that cannot be gathered in clinician-reported assessments (Wiklund, 2004). Additionally, patient-reported outcomes may allow discovery of patient concerns or conditions that have otherwise been missed and assist in the communication between patient and practitioner (Valderas et al., 2008).

Along with the advantages, there are some disadvantages to the use of patient-reported outcomes. For instance, some questionnaires are lengthy, which potentially extends the patient's visit. The clinician must also score the questionnaire and interpret the results, which also takes additional time. Another barrier to using patient-reported outcomes is selecting the most proper instrument. Numerous instruments exist for ailments throughout the body, which creates difficulty in selecting the most appropriate and valid instrument for the intended use. However, valid patient-reported outcomes are available and it is strongly suggested that the benefits of using patient-reported outcomes outweigh the burdens (Valderas et al., 2008).

Patient-reported outcomes can be further broken down into generic (global), disease-specific, region-specific, and patient-specific (Button & Pinney, 2004; Martin &

Irrgang, 2007). Generic instruments can be used for various populations and injuries or conditions. They are not specific to any single condition, injury, or population. An example of a commonly used generic instrument is the SF-36 (Ware & Sherbourne, 1992), or its shortened version, the SF-12 (Ware, Kosinski, & Keller, 1996). These instruments can be used with various populations to measure overall health status, but cannot accurately detect changes in health status due to specific injuries or conditions and should not be used exclusively for such purposes (Button & Pinney, 2004; Farrugia, Goldstein, & Petrisor, 2011; Martin & Irrgang, 2007). Disease-specific instruments are used to assess health status of patients suffering from a specific disease (Button & Pinney, 2004; Farrugia et al., 2011; Martin & Irrgang, 2007). The Ankle Osteoarthritis Scale was developed to measure disability and symptoms in patients with osteoarthritis of the ankle (Domsic & Saltzman, 1998). The results of disease-specific instruments are not generalizable to diseases or populations beyond those for which the instrument is intended (Martin & Irrgang, 2007). Region-specific instruments are broader than disease-specific instruments, but are not as comprehensive as generic instruments (Farrugia et al., 2011). They can be used to assess the effects of injuries or conditions on a specific area of the body. The Foot and Ankle Disability Index (FADI; Martin et al., 1999) is an example of a region-specific instrument. Patient-specific instruments are designed to more accurately assess the health status of a particular patient. Therefore, the information gathered from these instruments is not generalizable to other patients (Martin & Irrgang, 2007).

A modern version of patient-reported outcome is a computer adaptive test (CAT). Whereas most patient-reported outcome instruments are "static," meaning that there are a specified number of questions and the patient is required to respond to all of them, computer-adaptive tests are "dynamic" (Snyder et al., 2008), meaning that the questions presented to the patient depend upon the response to each subsequent question. Therefore, the CAT is able to self-adjust the questions to precisely measure the patient's ability. The self-adjusting feature allows there to be fewer questions necessary to precisely measure the patient's ability level and gives it the ability to obtain the desired end-point in fewer questions and in less time than a tradition static instrument, which decreases the demand placed upon the patient. Another advantage of CAT is that a large amount of items can be stored and used only when applicable according to the patient characteristics (Helbostad et al., 2009). Ideally, there will be CAT options available to assess patient-reported outcomes of all parts of the body and for various pathologies, but for now we must primarily rely on traditional, static instruments.

Instrument Development

Various self-reported instruments (often referred to as "surveys" or "questionnaires") are utilized throughout society to gather important information from consumers and/or patients (Slattery et al., 2011). Often, the responses to these instruments are used to create public policies, guide medical decision-making, among other societal decision-making, and therefore, must be developed properly (Slattery et al., 2011). These "survey" instruments are nearly identical to tests. On tests, examinees respond to questions (items) and receive a score based upon their responses. Judgments regarding the examinee's aptitude, or ability, are then made according to the results of the test. This is also the general concept applied to outcomes instruments. Decisions for future interventions are based upon the examinee's, or patient's, performance. Because the outcome of the instrument (test) can have great decision-making impact, failure to create a high-quality instrument can be detrimental to the patient in various realms of society (i.e. return to work, return to competition, insurance reimbursement, etc.).

Developing an outcomes instrument, however, is not a simple task. Many phases are involved, the first of which is determining the purpose of the instrument. Specific to health care outcome instruments, practitioners and researchers must decide must decide if an existing instrument will be modified or if an entirely new instrument will be created (Slattery et al., 2011). Once it is determined what the instrument should measure, the population for which the instrument is intended must be identified (American Education Research Association, America Psychological Association, & National Council on Measurement in Education, 1999). If an instrument is used for any purpose or population for which it was not intended, the instrument, and its results, can be deemed invalid (American Education Research Association et al., 1999; Slattery et al., 2011). Next, items can be created and responses and scoring can be determined (American Education Research Association et al., 1999). Several components can be considered when creating new items including patient input, clinical observation, theory, current literature, and expert input. If the researcher believes the existing instruments are insufficient, then the decision may be made to create new items to be added to an existing instrument.

When setting out to define or describe a construct that has not been previously defined, it is common for experts in the relevant field to be consulted. For instance, a group of concussion experts convened in Vienna, Australia to discuss and come to a consensus regarding the definition, recognition, assessment, and treatment of concussion (Aubry et al., 2002). Another technique in which expert opinion is gathered to reach a consensus on a topic is the Delphi method. The Delphi method was initially used by the RAND Corporation during the Cold War to estimate the number of A-bombs necessary to increase the US military capability. The purpose of the Delphi method is to reach expert consensus without convening of the experts. The experts are questioned individually in a repeated fashion by means of interview or questionnaire until consensus is reached or until they begin to diverge from consensus. Convening of the experts is avoided in order to eliminate the influence that some experts may have on others. It promotes the expression of individual thoughts and ideas (Dalkey & Helmer, 1963).

A short-coming of the Delphi method is that there are no clearly defined standards for proper implementation, which has led to many variations of the technique. Because of the various ways in which the Delphi method has been implemented, its rigor has been criticized (Hasson, Keeney, & McKenna, 2000; Keeney, Hasson, & McKenna, 2001). Hasson et al. (2000) recognized the criticisms of the Delphi technique and delineated the issues associated with the research method. First, they pointed out that the research problem must be carefully considered because not all research questions can be appropriately addressed using the Delphi method. They emphasized that other data collection techniques, as well as the logistics of the Delphi method, should be considered prior to adopting it as the primary means of data collection (Hasson et al., 2000).

Second, they emphasized that the process of the Delphi method must be well understood and pilot testing with a small sample should be carried out prior to data collection. Following the first round of interviews, which typically entails a series of open-ended questions; the data is reduced and analyzed and provided to the experts by means of a questionnaire. The data is once again reduced and given back to the experts for a third time. This process continues until consensus is reached, the experts begin to diverge from a consensus, or until the number of expert responses continues to decrease with each round (Hasson et al., 2000).

Third, expert selection and the term "expert" are somewhat subjective. What defines an individual as an expert to one researcher may be different to another. (Boulkedid, Abdoul, Loustau, Sibony, & Alberti, 2011) stated that the panel of experts should be representative of all stakeholders of the study topic. They also added that it is just as important to interview patients as it is to consult the experts (Boulkedid et al., 2011). Additionally, because the researcher is conscientiously choosing the experts, the sample is not random. Therefore, the consensus achieved from this method may be influenced by the researcher and the experts' opinions. There is also no consensus regarding the number of experts required to formulate the panel. Too few experts will limit the amount of data collected, which may introduce bias, while too many experts will create complexity in data reduction and analysis. (Hasson et al., 2000)

Another issue with the Delphi method is that there are no standards as to what defines a consensus (Boulkedid et al., 2011; Hasson et al., 2000; Holey, Feeley, Dixon, & Whittaker, 2007; Keeney et al., 2001). A range of percentages of agreement have been reported in the literature, while other authors feel that using numerical representations of qualitative data is improper (Hasson et al., 2000; Holey et al., 2007). Holey et al., (2007) sought to define a point at which data collection can be stopped during Delphi studies. They used email as their means of data collection. The participants were asked to indicate their level of agreement with a series of statements (5 point scale ranging from "strongly agree" to "strongly disagree") and to provide comments on each statement. For each round of questionnaires, the following were calculated: "percentage of response rates, percentages of each level of agreement for each statement to compensate for varying response rates, median, range, and their associated group ranking using the importance ratings, mean (Ware et al., 1996) and their associated group ranking using the importance ratings, and weighted Kappa (K) values to compare chance-eliminated agreement between rounds" (Holey et al., 2007). They noted, from previous authors' work, that consensus is synonymous with agreement and agreement can be determined via collected opinions, central tendency, and stability, or the tendency an individual's responses to be consistent from one round to the next (Holey et al., 2007). The results showed that mean, SD, range, and median can be used to determine consensus. The SD and range can be used to indicate convergence, or strength of the agreement. Additionally, the authors concluded that Kappa values can be used to measure stability of responses. Finally, the authors concluded that a decrease in the number of responses for each round indicates convergence (Holey et al., 2007). Although the Holey et al. (2007) analysis showed additional ways of determining consensus, it is only applicable to their method of data collection.

Other debated issues, as pointed out by Hasson et al. (2000) are if the number of opinions expressed by the experts per question should be limited, and if infrequently mentioned responses should be carried in to the second round of data collection. The authors state that the Delphi method has been administered both ways, for both issues. Whereas the authors feel that eliminating items from the first round to the second, they also point out that leaving all items in can create an abundance of data, which can potentially influence consensus (Hasson et al., 2000).

Anonymity is a concern with the Delphi method. Because the researcher must be able to track which experts have responded and which have not, anonymity is difficult to achieve. However, anonymity between experts is certainly achievable since they do not meet in a group setting. Additionally, the researcher can induce anonymity by keeping the experts' responses anonymous (Hasson et al., 2000; Keeney et al., 2001).

Although it is widely accepted that the results of each round should be reported separately, there is no standard for how the data should be represented. Delphi results have been presented in various ways (i.e., graphically, central tendencies, variance, and ranks). Additionally, not only must the data be reported appropriately, but it must also be interpreted properly. More specifically, consensus does not necessarily mean that the correct answer has been reached, but rather it represents the opinions of a group of experts. (Hasson et al., 2000)

Despite the limitations of the Delphi method, it has become widely used as a systematic method of data collection throughout the health care profession. Because evidence is not always available in sufficient quantities, the Delphi method is the best available means to obtain expert opinion and consensus upon which certain decisions can be made (Boulkedid et al., 2011). Recently, the Delphi method has been used in a number of ways including: as an initial step in the development of the International Classification of Patient Safety (Thomson et al., 2009), to develop a set of diagnostic protocol for elderly patients who experience dizziness (Maarsingh et al., 2009), to engage stakeholders in the development of a non-profit community-based organization evaluation (Geist, 2010), to determine which outcomes are most appropriate to measure in clinical trials (Sinha, Smyth, & Williamson, 2011), to determine important predictors of persistent shoulder pain (Vergouw, Heymans, de Vet, van der Windt, & van der Horst, 2011), and to develop a heat illness screening instrument (Eberman & Clearly, 2011).

In a recent review, a group of researchers analyzed the use of the Delphi method as a means of generating quality-indicators in healthcare (Boulkedid et al., 2011). They analyzed 80 articles published in 2009 or earlier which used the Delphi method to choose healthcare quality indicators. The results of their analyses revealed that most Delphi studies involved a modified Delphi approach, including face-to-face interviews. Overall, the review showed immense inconsistencies in the implementation, as well as the reporting, of the Delphi method employed. Many studies failed to report response rates, definition of consensus, and timeliness of feedback. Without knowledge of the procedures, the results of the Delphi method will be inappropriately interpreted

(Boulkedid et al., 2011). Therefore, it seems that if researchers increase the accuracy with which they report their methods, the usefulness of the Delphi method will be greatly enhanced. Many of the issues with the Delphi method could potentially be eliminated if researchers would precisely report their data collection methods. Although the Delphi is not without shortcomings, it shows value in obtaining group consensus.

As supported by Streiner and Norman (1995), expert opinion can be a valuable method of gathering information to assist item development. When experts are carefully selected, their opinions can offer the most up-to-date information in the topic of interest (Streiner & Norman, 1995). However, this technique also has shortcomings. There are no established standards as to the proper method of collecting expert opinion. It can be performed casually in conversation, or formally in an interview or meeting format. If the experts' opinions or areas of expertise are unbalanced and do not represent the full range of expertise, the data can be distorted towards one group's opinion (Streiner & Norman, 1995). In a recent attempt to develop a new instrument to assess the motor function in children with cerebral palsy, Wilson et al., (2011) used a small focus group consisting of three physical therapists to develop items to be added to an online survey. The focus group was intentionally small in order to facilitate attentive discussion (Wilson et al., 2011).

In addition to expert opinion, gathering patients' perceptions of how the injury or condition is impacting their life, their reasons for medical consultation, and their anticipated outcomes should be of utmost importance in the development of an outcomes instrument (Parker, Nester, Long, & Barrie, 2003). Input from patients or from people who have experienced the injury or condition to be investigated provides common themes that can be used to later formulate the new items. Once the new items are created, they are reviewed by the focus groups. The focus group participants then offer feedback regarding the items' understandability, applicability, clarity, and inclusiveness (Streiner & Norman, 1995). Although there are no commonly accepted standards for how many patients should be included, it has been suggested that a small number of participants is required in the item generation process (Bottomley, Jones, & Claassens, 2009). The participant characteristics, however, should reflect the characteristics of the intended population of the instrument (Bottomley et al., 2009).

Clinical observation can be one of the most productive means of item creation. Although it can be argued that theory should be taken into account prior to clinical observation, both often are used simultaneous. Both can be used to formulate instrument items based upon common ideas and observations. The ultimate shortcoming of integrating theory into the development of new items is that the theory may be inaccurate, which could subsequently lead to removal of one or more items from the instrument due to its imprecision. Therefore, research evidence is needed to support or refute theoretical constructs (Streiner & Norman, 1995).

Streiner and Norman (1995) emphasize that these methods of item development need not be used in isolation. Various combinations of item development techniques may be used as deemed appropriate by the investigator. In fact, they point out that it is atypical for only one method to be implemented in the creation of new instrument items (Streiner & Norman, 1995).

Once developed, the items must be evaluated for evidence of validity, reliability, responsiveness, readability, patient burden, alternative modes of administration, and it's potential to be altered to various cultures and languages (Valderas et al., 2008). Evidence of validity and reliability will be discussed in detail in the following segment of this literature review.

Dimensionality

The number of dimensions that the instrument will measure must be determined. For instance, an instrument may be used to assess quality of life, physical function, cognitive function, and psychosocial status (stress, anxiety, depression, etc.). Each of these could be considered a separate dimension or construct of overall health status. If multiple dimensions are to be used, factor analysis is necessary to group similar items to a common factor, which helps to further reduce the items needed on the instrument. However, it is ideal for an instrument to be unidimensional. In other words, all items on the instrument should load to the same factor, indicating that they are all measuring the same construct. Factors are constructs or theories that assist in the interpretation of the stability of a set of data. The goal of factor analysis is to use the least number of items necessary to explain the greatest amount of variance of a correlation analysis. Factor analysis, therefore, should be included in the analysis of multi-dimensional instruments because it allows interpretation of the characteristics of interest in the most efficient manner (Tinsley & Tinsley, 1987).

Measurement Properties

A critical issue in developing or selecting any outcomes instrument prior to its implementation in practice or research is identifying its measurement credentials or properties. The concept and process of identifying the measurement properties of an instrument is referred as psychometrics, and more recently clinimetrics. In the late 1800's, Sir Francis Galton began using psychometrics to test individual differences in height, weight, hand strength, among other variables. It is used to determine the validity of instruments which use several items to measure one construct (de Vet, Terwee, & Bouter, 2003). Since the 1800's, psychometric testing has been recognized as an integral component of instrument development. Also known as clinimetrics, psychometrics evolved as a means of assessing the validity of instruments which use one index to measure several constructs (de Vet et al., 2003). For example, clinimetrics are important in medicine when various symptoms result in one composite score, whereas in psychology psychometric analyses, such as factor analysis, are used to assess multiple constructs of a single dimension (de Vet et al., 2003). Fundamentally, psychometrics and clinimetrics serve the same purpose, but are used for different types of instruments. In fact, they are at times indistinguishable, as in the assessment of quality of life. Here, the boundaries between clinimetrics and psychometrics converge due to the nature of the assessment. Essentially, both clinimetrics and psychometrics are important evaluations of instrument validity (de Vet et al., 2003).

The most critical measurement property of any instrument or test is validity. Most of the measurement properties fall under the classification or validity or are secondary in importance. Validity is "the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests" (American Education Research Association et al., 1999, p. 9) and is therefore "the most fundamental consideration in developing and evaluating tests" (American Education Research Association et al., 1999, p. 9). Establishing evidence of validity involves both instrument construction and interpretation (American Education Research Association et al., 1999). Furthermore, when addressing the validity of an instrument it is imperative that multiple sources of validity evidence are accounted for. Evidence of validity should never be based on a single measure.

Construct Validity

The concept of construct validity has been debated in the literature. For instance, recently construct validity has been considered to be a unitary concept (T. Brown, 2010) which includes evidence of validity for test content, response processes, internal structure, relationships to other variables, and consequential aspects of construct validity. This is a departure from the former concept that construct validity was a component of overall validity. Finch, Brooks, Stratford, and Mayo (2002) described construct validity slightly differently. They stated that under circumstances in which no criterion standard has been established (i.e. pain), perhaps due to difficulty in measurement, construct validity is assessed (Finch et al., 2002). First, theories are formed in regards to the characteristic to be examined, then the instrument undergoing exploration is assessed to examine the extent to which its results support or refute the theories (Finch et al., 2002). If the results do not support the theories, the instrument lacks construct validity (Slattery et al., 2011). Construct validity has also been defined more specifically as convergent or divergent validity.

Convergent validity. Convergent validity exists when more than one instrument which assess the same construct are strongly correlated (Denegar et al., 2008).

Divergent validity. Instruments demonstrate divergent validity when they measure different constructs and are not strongly correlated (Denegar et al., 2008).

Face Validity

Face validity is the most simple, and least powerful, form of validity and implies that the instrument seems to measure what it is intended to measure based upon the items included in the instrument (Bellamy, 2005; Slattery et al., 2011).

Content Validity

If any aspects of the measure are left out of the assessment, then the instrument lacks content validity. It is the concept that all facets of the measure of interest are assessed in order to obtain a comprehensive depiction of the measure (Finch et al., 2002) and also ensures that the items fit the intent of the instrument (Slattery et al., 2011).

Criterion Validity

Criterion validity is the correlation of an instrument to a pre-existing gold standard measure of a criterion (Denegar et al., 2008; Finch et al., 2002). Criterion validity is concurrent or predictive.

Concurrent validity. The correlation between a new test and a pre-existing gold standard criterion assessment which are conducted simultaneously is known as concurrent validity (Denegar et al., 2008).

Predictive criterion validity. Predictive criterion validity indicates an instrument's ability to predict a criterion score from a gold standard which will be acquired in the future (Denegar et al., 2008).

There are several types of validity which should be considered when developing an instrument. It must be noted, however, that validity is a unified concept for which multiple pieces of evidence are needed. Validity should never be based upon a single piece of evidence (American Education Research Association et al., 1999).

Responsiveness

A valid instrument must be able to measure the patients' change over time. This is referred to as responsiveness and is a component of construct validity (Streiner & Norman, 1995). Although responsiveness is an important aspect of a patient reported outcome measure, it should be accompanied by a measurement of the minimally important difference (Pugia et al., 2001; Revicki, Hays, Cella, & Sloan, 2008). The MID is a measurement of the clinical impact of the change over time as based upon the minimal change necessary to cause patient-perceived benefit or harm (Revicki et al., 2008). Revicki and colleagues (2008) recommend that patient input be given utmost important in the measurement of MID. Because it remains unclear if MID refers to within-patient, within-group, or between-group change; patient input is important to determine the amount of change necessary to affect HRQOL. Additionally, the MID may vary from one patient to another depending on the severity, or state, of injury or illness (Bottomley et al., 2009). Several different methods can be used to determine the MID, which typically results in a range of acceptable MID. It is suggested that consultation with a group of experts and clinicians, via the Delphi method, be used to determine the MID of a particular patient reported outcome (Revicki et al., 2008).

Reliability

Reliability is the degree to which a measurement is repeatable or reproducible under various circumstances over time and produces the same results (Streiner & Norman, 1995). Although an instrument that is not reliable is not valid, reliability alone is not enough to deem an instrument valid (Cook & Beckman, 2006). An instrument must be reliable in order to be considered appropriate for use. Several types of reliability can be measured:

Test-retest reliability. Test-retest reliability is the concept that an instrument can be completed under the same or similar conditions over time and the results will be the same. It is an assessment of the steadiness of the patients' responses when no interventions are implemented (Finch et al., 2002).

Intrarater reliability. Intrarater reliability is the degree to which the results of an assessment or measurement performed by the same administrator are consistent over time (Streiner & Norman, 1995).

Interrater reliability. Interrater reliability refers to the ability of more than one individual to administer an assessment and produce the same result. It reduces the likelihood of error due to variations in implementation by the administrators of the instrument or assessment (Finch et al., 2002).

Internal consistency. Internal consistency is measured when multiple items are used to calculate a total score (Finch et al., 2002). If a group of items are used to produce the total score, the items should be correlated with one another Internal consistency is a measure of the correlation between items in a particular measure (Cook & Beckman, 2006). Cronbach's alpha has been commonly used to measure the internal consistency of the items which compose an instrument (Unick & Stone, 2010) and has been commonly incorporated in the process of validating instruments. However, a recent editorial published by Schweizer (2011) in the European Journal of Psychological Assessment questions the use of Cronbach's alpha. Schweizer claims that Cronbach's alpha is a true measure of consistency, but not of homogeneity, and is therefore, not the most appropriate measure to use when measuring particular characteristics or abilities (Schweizer, 2011). Instead, Schweizer proposed the use of the omega coefficient, which relies on loading of factors, rather than correlations. Thus, it is less likely to be influenced by external factors, which will decrease the likelihood of overestimation. Although its use has been criticized in some circumstances, Cronbach's alpha coefficient can serve as a starting point in item analysis.

Sensitivity

The sensitivity of an instrument is its ability to detect an injury or condition when the injury or condition truly exists (true positive; Streiner & Norman, 1995).

Specificity

Specificity is the ability of an instrument to correctly classify patients as not having an injury or condition (true negative; Streiner & Norman, 1995).

It is imperative that each of the aforementioned measurement properties is considered when creating a new instrument, or modifying an existing instrument.

Psychometric properties must be measured in order to determine the instrument's validity with a given population. If the instrument does not undergo appropriate analysis prior to use, its results are deemed invalid.

Determining Measurement Properties

Classical Test Theory

Classical Test Theory (CTT) has been the traditional method for instrument validation. The foundation of CTT is based on the assumption that there is a true score for each person's latent characteristic (Unick & Stone, 2010). CTT is based on the following model:

$$\text{Observed score} = \text{True score} + \text{Error}$$

CTT assumes that the true score is constant, any changes in the observed score are caused by error, and errors occur randomly and are not related to the true score or to each other (Panayides, Robinson, & Tymms, 2010).

This formula provides the framework for statistical analysis of CTT, such as Cronbach's alpha, along with other measures of reliability (Unick & Stone, 2010). Recently, however, Schweizer (2011), in the European Journal of Psychological Assessment, questioned the use of Cronbach's alpha. Schweizer is not alone in his critique of Cronbach's alpha or the overall use of CTT. Other authors have also pointed out the limitations of CTT and have suggested that other methods of instrument validation be used. For instance, it has been strongly recommended that researchers shift from relying on CTT as the primary means of assessing reliability and validity to incorporating item response theory (Michell, Ross, Blackburn, Hirth, & Guskiewicz, 2006; Panayides et al., 2010; Unick & Stone, 2010). IRT will be discussed in greater detail later in this review of literature.

A first limitation of CTT is that the difficulty of the items and the ability of the respondent (the latent characteristic) cannot be identified (Panayides et al., 2010; Unick & Stone, 2010). The credentials of the instrument are sample-dependent and test/item dependent, therefore, the difficulty of the instrument changes according to the ability level of the respondents (Denegar et al., 2008). Additionally, the items and examinees are measured on different scales, which inhibits proper comparisons of the items and respondents. The respondent is given one reliability score and one standard error of the measurement for the sum of the items on the instrument, rather than analyzing the responses to each item based upon each item's difficulty and the ability of the patient (Unick & Stone, 2010). Typically in classical test theory analyses, ordinal data is used inappropriately to calculate means or total scores. More specifically, statistical analyses performed with ordinal data are inappropriate because the analyses were designed for interval data.

However, there are advantages of using Classical Test Theory (CTT), which in some cases makes it more appropriate for use. A smaller sample is needed, as compared to IRT, to perform CTT analyses. This, of course, eases the burden of the investigator. CTT analyses are also simpler as compared to IRT and are generally easier to interpret. The assumptions of CTT are also typically satisfied by commonly used analyses (Schumacker, 2005). Because of the beneficial components of CTT, it is often used as a preliminary step in providing evidence of validity. The more complex IRT analyses are then conducted to gain more precise measures of validity evidence.

Item Response Theory

The basic assumption of item response theory (Michell et al., 2006) is that there is a certain probability that an individual with a certain ability level will respond correctly to certain items (Unick & Stone, 2010). The primary advantage of using IRT rather than CTT is that it is more precisely directed at the respondent's ability level. Therefore, it produces greater reliability and lower standard error of the measure because the items are directed specifically at the patient's ability level. Additionally, IRT can provide different reliability and standard error of the measure values for people of different ability levels (Unick & Stone, 2010). For instance, ankle injuries occur on a spectrum, from mild to severe. An instrument whose items only assess simple tasks, such as transitioning from the seated to standing position or walking on a level surface, gives only minimal insights on the ability of the mildly injured patient. On the other hand, an instrument that contains highly difficult items, such as sprinting or cutting, provides very little information regarding the ability of severely injured patients. In other words, the more closely matched the items are to the ability level of the respondent, the more accurate the measure will be (Unick & Stone, 2010).

The initial step when using IRT is to transform the data to logits, which allows interpretation of the patient's ability level independently of the items, or to a comparison group or normative data (Baylor et al., 2011). The higher the logit value, the higher the probability the patient will respond to the item correctly, indicating higher patient ability

(Baylor et al., 2011). Typically, the logic scale ranges from -3.0 to 3.0 logits, where a logit of 0 represents the mean ability of the test group, or the mean of the difficulty level of the item (Baylor et al., 2011). Once the raw data is transferred to logits, then classical statistical analyses can be appropriately performed (i.e. *t*-tests, correlations) because the ordinal data is transferred to interval data.

The precise measurement capabilities of IRT, unlike CTT, allow it to accurately measure change over time (Unick & Stone, 2010), which is essential when assessing the effectiveness of a treatment intervention. Additionally, IRT is able to obtain strong reliability measures while using fewer items, which is beneficial when measuring more than one characteristic. The reduced number of items also eases the burden of the respondent (Unick & Stone, 2010).

Rasch Model

The greatest separation between CTT and IRT is that IRT is based upon mathematical models. The Rasch model is the simplest mathematical IRT model as it contains only one item parameter (item difficulty) and was designed for use with dichotomous items (Baylor et al., 2011; Thomas, 2011). It was developed in 1960 by Georg Rasch, a Danish mathematician and was primarily designed for reading tests. Ben Wright and colleagues from the Measurement, Evaluation, Statistics, and Assessment (MESA) Unit from the University of Chicago, worked to broaden its applicability to other areas of education. Others involved in the diffusion of the Rasch model to additional disciplines were Bruce Choppin, David Andrich, and Geoff Masters. Whereas

Rasch himself recognized that the model did not fit all areas, these men transitioned the model for use outside of the domains for which it was created (Panayides et al., 2010).

The three variables included in the Rasch model include ability of the patient, difficulty of the items, and the probability that the patient will respond correctly to the item. The items are graphically plotted to display the item characteristics. The *x*-axis represents difficulty of the item as well as patient ability. The *y*-axis represents the probability of the patient to respond correctly to the item (Baylor et al., 2011). Typically, computer-based programs are used to conduct these analyses.

The benefits of the Rasch model have been cited by (Unick & Stone, 2010). The first is that it measures the patient's ability and item difficulty on the same metric, allowing direct comparison of the patient's ability and the item. Secondly, the Rasch model is able to calculate the difference between the estimated item responses and the actual item responses, known as "item fit," which allows selection of items that appear to fit the model and prompts further investigation of misfit items (Unick & Stone, 2010). Essentially, the Rasch model can provide a more accurate assessment of the patient's ability level, which in turn assists the clinician in selecting items that are appropriate for the patient's ability level.

In June 2009, the Lower Extremity Functional Scale (LEFS) was evaluated using the Rasch model (Lin, Moseley, Refshauge, & Bundy, 2009). The instrument was developed to assess the activity level of patients with lower extremity musculoskeletal pathology, but the purpose of this study was to assess its clinimetric qualities in patients with ankle fracture. Internal consistency and construct validity were evaluated by transforming the raw data to interval scores using logits. The internal consistency of the instrument was calculated using the internal reliability coefficient of the scale, which is similar in nature to Cronbach's alpha, and by correlating each item to the overall scale. Construct validity was determined by taking the variance in activity limitation explained by the LEFS and analyzing it using principal components analysis of the residuals (Lin et al., 2009). A hierarchy for each item and for each participant was developed. The difficulty of the items and the ability of the participants were compared with the assumption of the Rasch model that states that items of lesser difficulty will be responded to correctly by all participants and items of greater difficulty will be responded to correctly by participants of higher ability.

The comparison was depicted using two pairs of infit and outfit statistics, which represent how well the items fit the model. Infit statistics measure responses to items that are similar to the participant's ability, whereas outfit statistics measure responses to items that are dissimilar to the participant's ability. Concurrent validity was examined at 4 to 6 weeks (short-term) follow-up and at 24 to 26 weeks (medium-term) follow-up. Pearson correlation coefficients were used to measure relationships between the LEFS and walking speed, stair-stepping rate, and the Olerud Molander Ankle Score again at the short-term and medium-term follow-up. Responsiveness was also measured at both the short-term and medium-term follow-up using effect size and standardized response mean. A high internal responsiveness score was set at 0.8 or higher. External responsiveness was measured using Guyatt responsiveness ratio and the receiver operating characteristic curve. Floor and ceiling effects of the LEFS total score were measured at baseline, short-term, and medium-term follow-up. If more than 15% of the participants obtained the lowest score, a floor effect existed. If more than 15% of the participants received the highest score, a ceiling effect existed (Lin et al., 2009).

The Rasch analysis revealed that the LEFS had high internal consistency in participants with ankle fracture at all measurement times. Item correlation was also similar at all three measurement time points. All items on the instruments were related to the overall scale, as shown through positive correlations between the items and the scale. Most of the items fit the scale. Concurrent validity of the LEFS was superior as compared to Olerud Molander Ankle Score at the short-term and medium-term follow-ups. The internal responsiveness of the LEFS was high at the short-term and medium-term follow-ups, and external responsiveness was high at short-term follow-up. The Guyatt responsiveness ratio was low at medium-term follow-up, which is undesirable. The LEFS did not show a floor-effect, but came close to showing a ceiling effect with 14% of participants achieving the highest possible score at the medium-term follow-up (the tolerance level was set at 15%; Lin et al., 2009).

Multiple methods are available to determine the psychometric properties of measurement instruments; it is the duty of the investigator to decide which method is the most appropriate for their intended use. Although Classical Test Theory has been criticized, it seems to have merit for certain research purposes, and offers a baseline for which further, more complicated analyses can be conducted.

Instrument Selection

Selecting the most appropriate outcome instrument for the intended use is of utmost importance to the proper interpretation and applicability of the instrument. Each of the aforementioned components should be present in a truly valid instrument. Unfortunately, however, few instruments which assess the outcomes of foot and ankle injuries have undergone in-depth psychometric testing. In addition to an instrument's psychometric testing and analysis, other factors must be considered when choosing an instrument. This section of the review of literature will address how to select the proper instrument according to the specific characteristics of the patient for which it will be used.

An important component of an outcomes instrument is its ability to discriminate patients of various ability levels. In order to differentiate between patients both at the high and low ability levels, and have the capability to measure change over time, the instrument should be free of ceiling and floor effects. A ceiling effect occurs when an instrument is not difficult enough to detect the activity limitations or functional improvements in high ability patients. The patients are able to achieve a score of 100% on the instrument, which typically indicates that there are no functional limitations, when in fact; the patient is not able to participate at the desired level of function. The instrument provides a false impression that the patient is fully recovered. If an instrument is to accurately assess the function of high ability individuals, it must not have a ceiling effect. Conversely, an instrument must be able to measure worsening of a condition or injury. If it cannot do this, it is said to have a floor effect. In other words, patients will continue to receive the lowest overall score possible, even though their condition or injury is declining. This provides inaccurate information about the changes in the patient's condition and deems the instrument inappropriate for use (Finch et al., 2002).

Valderas and colleagues (2008) developed a tool designed to assist instrument selection. The tool, known as EMPRO (Evaluating the Measurement of Patient-Reported Outcomes), contains 39 items which assess eight components of a patient-reported outcome. The eight components include conceptual and measurement model, reliability, validity, responsiveness, interpretability, burden, alternative modes of administration, and cross-cultural and linguistic adaptations (Valderas et al., 2008). The authors stress that each of these components should be evaluated when choosing an outcomes instrument.

Instrument Modification

In some cases, an appropriate patient-reported outcome instrument may not exist, depending on the patient characteristics. In these situations, it may become necessary to modify an existing instrument, rather than create an entirely new one, as time is often of the essence. It is then necessary to determine how the instrument will be modified. Typically, modifications that may be made include changes in wording and/or content, instrument administration, language or cultural revision, or transition to a new population (Snyder et al., 2008).

When it is deemed necessary to modify an instrument, the following criteria must be upheld: (1) the instrument of choice has been assessed for evidence of validity, (2) the newly revised items are applicable and suitable for the instrument's intended use, (3) the instrument itself is appropriate for use with the intended population, and (4) new instructions on how to properly interpret the results of the instrument are provided (Snyder et al., 2008).

A complete re-validation of the newly created instrument may not be necessary (American Education Research Association et al., 1999), but the requirements range from minor revalidation necessary to complete revalidation (Snyder et al., 2008). The requirements for revalidation depend upon the amount of changes that were made. Minor changes in wording, content, instructions, or response choices typically do not warrant revalidation of the entire instrument (Snyder et al., 2008). As opposed to a common misconception, when generic outcome instruments are applied to patients with specific injuries or diseases, revalidation should not be necessary, as long as the instrument has undergone rigorous psychometric testing (Snyder et al., 2008). Similarly, when disease-specific instruments are used to measure outcomes of patients with an injury or disease that is characteristically similar to the disease for which the instrument was developed, no revalidation should be warranted. On the other hand, if a disease-specific instrument is applied to patients whose disease or injury is vastly different than the disease for which the instrument was developed, or if considerable changes in content are made, then complete revalidation of the instrument may be necessary (Snyder et al., 2008).

Although there are many steps involved in instrument development, testing, and selection, it is crucial that appropriate and psychometrically sound instruments are used in research and clinical practice. There is not a more vivid example of the need for appropriate outcome instruments than in the assessment and rehabilitation of ankle injury.

Understanding Ankle Function

At the root of addressing outcomes is measuring patient function. Function has been defined as "the physiological or psychological functions of body systems" (World Health Organization, 2002). According to the WHO ICF, body functions and structures include the entire human body and psyche. Therefore, when assessing outcomes following injury, the patient's entire being must be considered, not just the injured part. Furthermore, the human body functions mechanically in a series of "links," holistically known as the kinetic chain. During weight bearing activities for example, each joint and body segment (i.e., foot, ankle, knee, pelvis, spine, shoulder, etc.) function as a collective unit rather than an isolated part to achieve function. Therefore, pathology at the foot or ankle can impact overall function, even though the other links, such as the hip and spine, are void of pathology. Due to the ankle's role in the kinetic chain, injury can result in vast functional limitation. Any activity requiring ambulation or weight-bearing will be hindered because the ankle plays an important role in the body's ability to function during weight-bearing activities. Therefore, it is vital for clinicians to accurately monitor and assess injury outcomes.

The need for precise ankle outcomes assessment is clearly depicted by the fact that ankle sprains, specifically lateral ankle sprains (LAS), represent the single most common athletic injury (Hootman et al., 2007) and the strongest predictor of sustaining an ankle sprain is history of a previous ankle sprain (McKay, Goldie, Payne, & Oakes, 2001). In an epidemiological study of high school athletic injuries, Fernandez, Yard, and Comstock (2007) found that ankle injuries had the highest incidence rate for both boys and girls. Similarly, Borowksi et al. (2008) found that ankle and foot injuries were the most commonly sustained injury in high school basketball players and Sankey et al., (2008) reported lateral ankle ligament injuries to be the most common in professional rugby. Furthermore, as many as 70% of people who sustain an ankle injury will experience repetitive sprains or chronic injury symptoms (McKay et al., 2001). It would, therefore, seem logical that the field of athletic training would have firmly established outcomes assessment procedures for ankle injuries. Unfortunately, this is far from the truth. The startling truth is that outcomes which substantiate treatment of ankle injuries are limited. One barrier is the availability of valid outcome instruments; of the more than 40 instruments used for the ankle and foot (Button & Pinney, 2004), none have been validated for athletic training clientele. This potentially creates an issue when these instruments are used to assess patients with high levels of function, as seen in athletes. The athletes appear healthy, even though they may still have functional limitations which may not be detected if the items on the instruments are too easy; that is, they fail to address the higher ability levels of physically active individuals.

Further compounding the problem of assessing outcomes following LAS is their repetitive nature, which has led to the identification of a phenomenon called "chronic" or "functional" ankle instability (CAI, FAI, respectively; Freeman et al., 1965; Hertel, 2002; Tropp, 2002). Although a common entity in practice and research, FAI is inconsistently defined and assessed (Howe-Weichers, 2007). The first inconsistency regarding FAI is its definition. Despite widespread assimilation, and overwhelming support for its existence, there is currently no consensus regarding its definition. To further complicate the issue, authors inconsistently use the terms "functional," "chronic," and "lateral" ankle instability, or just "ankle instability."

While some authors feel that FAI and CAI are two terms which describe the same entity, others feel that they are two different entities. The original definition of functional ankle instability was coined by Freeman (1965). He initially described functional ankle instability as the tendency for the foot to "give way" (Freeman, 1965, p. 669). Later, functional instability was described as "the disabling loss of reliable static and dynamic support of a joint" (Vaes, Duquet, Casteleyn, Handelbert, & Opdecam, 1998, p. 692). Additionally, lateral ankle instability has been described as "the existence of an unstable ankle due to lateral ligamentous damage caused by excessive supination or inversion of the rearfoot" (Hertel, 2002, p. 364) and chronic ankle instability as "the occurrence of repetitive bouts of lateral ankle instability, resulting in numerous ankle sprains" (Hertel, 2002, p. 364). In his research, Hertel does not use the term "functional ankle instability," but instead supports a paradigm in which mechanical and functional insufficiencies as sub-components of chronic ankle instability. Tropp (2002) also supports the concept that functional and mechanical instabilities are components of chronic ankle instability, but describes functional ankle instability as "the subjective feeling of ankle instability or recurrent, symptomatic ankle sprains (or both) due to proprioceptive and neuromuscular deficits" (Tropp, 2002, p. 512). Interestingly, Tropp's definition of FAI is very similar to Hertel's definition of CAI. This further demonstrates an inconsistency in the terminology used throughout the literature.

The term "chronic lateral ankle instability" has been described as giving out of the ankle, typically accompanied by at least two or three previous severe lateral ankle sprains (DiGiovanni, Partal, & Baumhauer, 2004). In 2010, a group of authors added pain and swelling to the more traditional definition of CAI (recurrent sprain, mechanical instability, giving way, functional instability; Delahunt et al., 2010). Recently, Hiller, Kilbreather, and Refshauge (2011) proposed a new model of CAI that was derived from the Hertel (2002) model. They took the two subgroups of CAI proposed by Hertel (2002) and added a third, recurrent sprain. Furthermore, they pointed out that the subgroups could co-exist, or exist independently, resulting in a total of seven potential subgroups.

Not only are there various definitions used to describe CAI and FAI, but the use of ambiguous terms throughout the various definitions has led to even further confusion. For instance, it is often unclear how many sprains deem the injury to be "recurrent," the number of sprains and how far in the past defines "history of previous sprain," what defines "significant" injury, and how many times or how often the ankle must "give way." None of these components are consistently defined and in some cases, are not defined at all. Another area of debate and disagreement in terms of CAI is the presence of ligamentous laxity. While some researchers believe that ligament laxity is a contributing factor to CAI (Hubbard, 2008) others state that it is incorrect to assume that ligament laxity is present in all cases of CAI (Denegar & Miller, 2002).

Yet another inconsistency in regards to FAI and CAI is that there is no consensus regarding the identification of characteristics present for those who have these phenomena. Researchers have classified the existence of FAI and CAI inconsistently and clinicians have diagnosed it arbitrarily. To substantiate the existence of any injury or condition, it must be definable and have a measurable impact on function. Neither of these is true regarding FAI or CAI. To demonstrate the vast disparity in identifying individuals with FAI or CAI, below is a list of various inclusion criteria that have been used:

1. "(1) history of at least one significant lateral ankle sprain in which the subject was unable to bear weight or was placed on crutches, (2) episodes of at least one repeated lateral ankle injury or feelings of ankle instability or "giving way," and (3) no present participation in a rehabilitation program" (Demeritt, Shultz, Docherty, Gansneder, & Perrin, 2002, p. 508).
2. "history of mild to moderate ankle sprain at least 12 months before the study that required immobilization or non-weightbearing status for three days" (C. Brown, 2011, p. 2)
3. "(1) self-report of a history of one sprain followed by at least three days of immobilization and (2) self-report of at least two ankle sprains and at least two episodes of "giving way" sensations during physical activity within the year before the subjects' enrollment in this study" (Ross, Guskiewicz, Gross, & Yu, 2009, p. 400).
4. "previous history of unilateral ankle sprain, frequent giving way of the ankle (at least once a month), pain, feeling of instability, and decreased function" (Hubbard, Kramer, Denegar, & Hertel, 2007b, p. 362).

5. "at least three episodes of lateral ankle sprain or instability with ongoing symptoms for at least six months" (Van Bergeyk, Younger, & Carson, 2001, p. 38).
6. "history of at least one significant unilateral inversion sprain of either ankle, followed by episodes of at least one repeated injury or feeling of ankle instability or giving way, no symptoms of acute injury at the time of testing and no reported history of surgery or fracture...free from mechanical instability as tested by anterior drawer and talar tilt tests" (Hadadi et al., 2010, p. 2) .
7. "history of FAI with no formal rehabilitation" (McKnight & Armstrong, 1997, p. 22).
8. "history of more than one ankle sprain, with the original injury occurring at least 12 months prior, and residual symptoms, as quantified by four or more *yes* responses on the Ankle Instability Instrument, self-reported symptoms of disability due to ankle sprains of 90% or less on the FADI and FADI-Sport" (Knapp, Lee, Chinn, Saliba, & Hertel, 2011, p. 258).

Although some components of the inclusion criteria are the same between studies, no two studies use identical criteria. Clearly, there is a great deal of variation in the definition of FAI and CAI, a lack of agreement as to which term should be used, and great disparity in the inclusion criteria. These inconsistencies have created discrepancies in research results because no two groups of researchers classify the condition using the same criteria (Howe-Weichers, 2007). Therefore, cross-study comparisons are impossible and the true impact of the condition is left largely unknown. Without consistent classification, outcomes cannot be accurately measured. The lack of standard outcomes and the inconsistencies in identification of repetitive LAS has further compounded the reporting of ankle outcomes. Therefore, athletic trainers are neglecting to address one of the most important issues identified in their profession (outcomes) for the most common injury for which they provide care (LAS).

Foot and Ankle Outcomes Instruments

Selecting an outcome instrument for foot and ankle injuries can be challenging for clinicians. Over 40 instruments exist to assess foot and ankle outcomes following injury (Button & Pinney, 2004). However, there is no consistently used outcomes instrument to determine if the patient is fully recovered from injury. Further complicating clinicians' decisions is the lack of psychometric testing of many of the available instruments. Specifically, there is no instrument that has been validated for assessing the functional capabilities of individuals with high levels of ability. The tasks assessed with the available instruments are typically directed towards activities of daily living. There is a need for a valid instrument which can accurately assess the functional status and outcomes of people across the entire physical ability spectrum. Another issue is that clinicians must be able to select the most appropriate instrument for each case and have the ability to correctly interpret the score (Martin & Irrgang, 2007). The instruments that are most commonly used throughout the literature for foot/ankle outcomes and appear to have the most potential for athletic training clientele based on item relevance include: the

Foot Function Index (Chung, Ma, Lau, & Griffiths, 2012), Ankle Function Score (AFS), American Osteopathic Foot and Ankle Score (AOFAS), Foot and Ankle Disability Index (FADI), and the Foot and Ankle Ability Measure (FAAM). This portion of the literature review will delineate the most commonly used self-reported ankle outcomes instruments, the basic components of function that they attempt to measure, how they were developed, and the current status of their psychometric properties.

Foot Function Index (FFI)

In the late 1980's, Budiman-Mak, Conrad, and Roach (1991) acknowledged that no instruments existed which assessed pain and joint pathology exclusively at the foot; all of the available instruments examined the whole body. Hence, they set out to create an instrument which could evaluate the effects of joint pathology specifically at the foot. In particular, they were interested in examining the patient's perceived pain, disability, and activity restriction (Budiman-Mak et al., 1991). In general, the FFI was created to assess change in the patient's state overtime, as well as current status. It was designed to be brief and simple because it was intended for use with the elderly population. It contained 23 items, which were categorized into three subscales: pain, disability, and activity limitation, as related directly to the foot. The participants rated each item on visual analog scales which were placed on horizontal lines. The extremes, ranging from "no pain" to "worst pain imaginable," "none of the time" to "all of the time" and "no difficulty" to "so difficult unable," were placed at each end of the line. The participants were asked to draw an indicator of their perceived pain or disability appropriately to reflect their experience in the last week. The pain subscale contained nine items, as did the disability subscale, while the activity limitation subscale consisted of five items. The items were scored by breaking the line into 10 equal parts and assigning each segment a score of 0 to 9. The subscale scores were derived by taking the total points for all items in the subscale, dividing them by the total points possible for the subscale, and multiplying by 100. The total FFI score was calculated by averaging the scores of the three subscales. A higher score indicated greater disability (Budiman-Mak et al., 1991). Although the authors reported how the instrument is scored, they did not report how the items were created.

Participants (n = 87) with either definite or classical active rheumatoid arthritis were recruited from multiple hospitals to participate in this study. The mean age of the participants was 61 (range = 24 – 79) and 89% of the participants were male (Budiman-Mak et al., 1991). The participants were placed into a treatment or control group and completed the FFI every six months for three years. Consistency of the instrument was calculated using Cronbach's alpha for each subscale as well as the total score. Construct validity was measured using principal components factor analysis to identify the non-correlated factors, as well as a varimax rotation to identify the correlated factors. Criterion validity was assessed by analyzing the correlations between the FFI total score and sub-scale scores, and painful foot joint count, time to walk 50 feet, and hand grip strength. The FFI's capability to measure change over time was measured by analyzing changes in an objective measure of the disease activity, the number of painful joints in the foot, and the changes in the FFI total score, as well as the subscale scores, during a six-month period of time (Budiman-Mak et al., 1991).

The test-retest reliability of the total FFI was the strongest (0.87), followed by the disability subscale (0.84) and activity limitation (0.81). The test-reliability of the pain subscale was moderate (0.695). The Cronbach's alpha analysis showed strong internal consistency of the FFI total (alpha = 0.9556), disability subscale (alpha = 0.9460), and pain subscale (alpha = 0.9276). The internal consistency of the activity limitation scale was moderate (alpha = 0.7327). The principal components factor analysis identified four factors. All pain subscale items loaded the most strongly on factor one, all disability subscale items loaded most strongly on factor two, three activity limitation items loaded on factor three, and the residual two items loaded on factor four. The relationship between the FFI total and foot pain joint count was moderate ($r = 0.5290$), as was the relationship between the FFI total and the 50 feet walking time ($r = 0.4785$).

The 50 feet walking time test was significantly correlated with all three subscales. The number of painful foot joints was strongly correlated with all three subscales as well, but the relationships were not as strong as with the 50 feet walk test. The painful hand joint count, as well as the grip strength measurement, had the weakest correlations with the subscales. However, there was a moderate relationship between grip strength and the disability and activity limitation subscales. Change over the six month time period in the number of painful joints in the foot showed a moderate correlation with FFI total change and change in pain scores. Painful joint count in the foot correlated significantly, but weakly, with activity limitation change scores. Painful joint count in the foot had no relationship with disability score changes (Budiman-Mak et al., 1991).

The FFI has undergone adequate psychometric testing in an elderly population with rheumatoid arthritis and was shown to have acceptable internal consistency, construct validity, criterion validity and is able to detect change over time (Budiman-Mak et al., 1991). It was later assessed for reliability in a group of patients with foot pathology who did not have systemic disease. It was discovered that the FFI is an adequate instrument for use with low functioning individuals because many of the patients' abilities exceeded the items on the instrument, indicating that the instrument has a ceiling effect (Agel et al., 2005). In 2006, the validity of the FFI was measured by correlating it with the Medical Outcomes Study Short Form-36 (SF-36). This analysis revealed moderate to high correlations between the FFI and the SF-36, which the authors deemed as adequate (SooHoo, Samimi, Vyas, & Botzler, 2006). The FFI was used by (Ibrahim et al., 2007) as a means of comparison with the subjective portion AOFAS rating scales. Although the FFI has been deemed as adequate for use in patients with foot and ankle pathology, it still does not contribute to the missing piece of the puzzle; it is suitable for the elderly population and individuals with rheumatoid arthritis, but the results of these analyses cannot be generalized to the high ability or healthy populations.

In 2002, a new instrument was created based upon the items contained in the FFI. Whereas the FFI is used to assess foot pathology related to rheumatism, the Ankle Osteoarthritis Scale was developed to assess patient reported restrictions in function due to osteoarthritis of the ankle (Domsic & Saltzman, 1998). The original FFI consisted of 23 items, separated into three subscales, whereas the AOS contained a total of 18 items, divided into two subscales (pain and disability). The items included in the pain and disability subscales of the FFI were retained for the same subscales of the AOS. The five items in the activity limitation subscale of the FFI were eliminated, resulting in 18 total items of the AOS. Validity and reliability testing of the AOS implied that it is an effect tool to evaluate the outcomes of patients with ankle arthritis (Domsic & Saltzman, 1998).

Ankle Function Score (AFS)

The ankle function score was initially developed in attempt to differentiate mild and severe acute ankle injuries (De Bie, De Vet, van den Wildenberg, Lenssen, & Knipschild, 1997). It was developed based on the Lysholm score for the knee. In particular, the AFS assesses pain, perceived instability, weight bearing, swelling, and gait pattern. To test the instrument, thirty-five patients were diagnosed with acute ankle injury in a hospital emergency room and were re-examined two and four weeks following the initial assessment. A bivariate analysis, as well as logistical modeling, was used to predict outcomes at two and four weeks post injury. The sensitivity for prognosis of recovery two weeks post injury was 97%. The specificity for the same time period was 100%. The authors concluded that the AFS can provide a fairly accurate injury prognosis following acute ankle injury and can differentiate between mild and severe injuries (De Bie et al., 1997).

A later study (Van Der Wees et al., 2010) sought to determine the prognostic validity, concurrent validity, and responsiveness of the AFS. Physiotherapists ($n = 20$) were recruited to assist in the data collection. The therapists collected AFS, Olerud and Molander Ankle Score (OMAS), and patient-specific complaints (De Bie et al., 1997) upon initial evaluation and every two weeks following. Global Perceived Effect (GPE) was collected two weeks following the initial evaluation and at the end of the treatment period (Van Der Wees et al., 2010). The OMAS was initially created as a self-reported assessment of symptoms following ankle fracture (Olerud & Molander, 1984). Its evidence of validity was established in 2004 (Haywood, Hargreaves, & Lamb, 2004). The PCS was created to monitor changes in primary complaints over time in low back pain patients by (Beurskens, de Vet, & Koke, 1996). This instrument has been shown to be responsive for reporting main complaints. GPE is a global rating scale used to measure self-reported change following intervention. It consists of a seven-point scale in which 1 indicates complete recovery, whereas a score of 7 indicates worse than ever (Van Der Wees et al., 2010).

Prognostic validity was established by comparing the AFS score to the GPE score during the initial evaluation. It was hypothesized that an AFS score of ≤ 40 at ≤ 5 days post-injury result in a GPE score of ≥ 2 at 2 weeks post-injury. Sensitivity and specificity were predicted to be 60 – 80%. The results indicated that the sensitivity and specificity of the AFS at baseline was 76%, and 63% 2 weeks post-injury, with GPE as the external criterion. The ability of the AFS to predict that patients with a severe injury would not recover in two weeks or less was 86%, while its ability to predict that patients with mild injuries would recover in two weeks or less was 45% (Van Der Wees et al., 2010).

Concurrent validity was assessed by comparing AFS scores with OMAS and PSC scores. Correlations were measured by comparing scores at initial evaluation, end of the treatment, and the change between initial and ending scores. Pearson's r for linear correlation was used for the AFS and the OMAS, and Spearman's ρ for non-linear correlation was used for AFS and PSC comparison. The Pearson's correlation between the AFS and OMAS at initial evaluation was 0.82, at end of treatment was 0.70, and total change was 0.79. The Spearman's correlation between AFS and PSC at initial evaluation, end of treatment, and total change were 0.26, 0.49, and 0.37, respectively (Van Der Wees et al., 2010).

Responsiveness of the AFS was examined using effective size, standardized response mean, and responsiveness ratio of Guyatt (GRI: average change of recovered patients/SD of average change of non-recovered patients). The effect size (2.00) and standardized response mean (2.10) of the AFS were similar to that of the OMAS (2.08 and 2.23, respectively). The standardized response mean of the AFS (2.10) was higher than that of the PSC (1.71), and the GRI of the AFS (2.15) was similar to that of the OMAS (2.39) and PSC (2.13; Van Der Wees et al., 2010).

The overall results of this study found limited evidence for the use of this instrument. However, the authors supported its continued use due to its ease of use and convenience (Van Der Wees et al., 2010). This advice should be interpreted with caution as there is no empirical evidence to validate the scores provided by this instrument. The authors essentially suggest that it is better to use a faulty instrument than to use nothing at all. Perhaps this perception is what has led to the development and use of an abundance of instruments for the foot and ankle, even though there is little or no evidence to support their use.

In a recent publication by (van Rijn, Willemsen, Verhagen, Koes, & Bierma-Zeinstra, 2011), the Ankle Function Score was used to assess how its score, along with pain intensity and giving way of the ankle, were related to self-reported recovery and how initial scores influence the relationship following acute lateral ankle sprain. The results of this analysis must be interpreted with caution as the AFS has not been previously validated. Even though the instrument was used properly, its scores cannot be generalized or practically applied.

American Osteopathic Foot and Ankle Score (AOFAS)

Originally published in 1994 by Kitaoka, et al., the American Orthopaedic Foot and Ankle Society rating scales contain four separate systems of rating function throughout the foot and ankle. The scales include the Ankle-Hindfoot Scale, Midfoot Scale, Hallux Metatarsophalangeal-Interphalangeal Scale, and the Lesser Metatarsophalangeal-Interphalangeal Scale. The development of the scales was based on previously established rating systems of the ankle, foot, knee, hip, spine, shoulder, elbow, wrist, and hand. A group consensus of the AOFAS determined that four new scales would be created, each of which had a maximum score of 100 points. Each item is given at least three choices, each of which is assigned a score. The patient is asked to indicate the most appropriate choice for each item. The total score is calculated by taking the attained score, dividing by the maximum possible score, and multiplying by 100. The scales were designed to be user-friendly, so the use of high-technology or cumbersome equipment was not included. Each scale involved both self-reported and clinician-measured items and were designed to be usable for various pathologies of the body part of interest in general population patients. In particular, the AOFAS assesses pain, activity limitation (ADL, recreation), and walking (distance, surface; Kitaoka et al., 1994).

Due to their clinician-reported components, the AOFAS instruments are not purely patient-reported. However, a validity and reliability study of the self-reported components alone was published in 2007 (Ibrahim et al., 2007). The FFI was used as the comparison scale in this study, as its validity and reliability had been previously established. The participants in the study were separated into two groups. Both groups were pre-surgical for ankle or foot operations. The first group ($n = 45$) responded to the AOFAS self-report component and the FFI prior to surgery and at least two weeks following surgery. The second group ($n = 43$), like the first, responded to the AOFAS self-report component and the FFI prior to surgery, but their second response was not gathered until three months following surgery. Pearson correlation coefficient was used to analyze the relationship between the AOFAS and FFI. Test-retest reliability and responsiveness were calculated for the second group using a paired *t* test to compare pre-surgical and post-surgical AOFAS scores. The Pearson correlation coefficient analyses showed no significant relationships between the AOFAS and the FFI. The *t* test used to measure test-retest reliability showed no significant differences between the scores. The *t* test used to analyze the pre- to post-operative scores of the AOFAS was significantly different, indicating that the instrument is able to detect change in self-reported quality of life. The authors concluded that the AOFAS scales produce reliable and valid scores as related to quality of life in patients with foot and ankle pathology (Ibrahim et al., 2007).

Interestingly, even prior to the Ibrahim et al. (2007) examination of the AOFAS scales, these scales were commonly used to assess patient-reported outcomes following various surgical procedures. For instance, Scranton, McDermott, and Rogers (2000) used the AOFAS Hindfoot scale to assess patient-reported outcomes following Brostrom surgical procedure due to ankle instability. Later, in 2005, another group reported their use of the AOFAS Ankle-Hindfoot scale to measure patient reported outcomes after undergoing the Brostrom-Gould surgical technique (Brodysky, O'Malley, Bohne, Deland, & Kennedy, 2005). In a study published in 2006, SooHoo et al. recognized that the AOFAS lacked proper evidence of validity testing, but used it in their study anyway because it is one of the most often used instruments to assess patient-reported outcomes for patients with foot and ankle pathology. Thus, the AOFAS are not the most ideal instruments to use.

Foot and Ankle Disability Index (FADI)

The Foot and Ankle Disability Index (FADI-Total) was originally developed by Martin et al., (1999) to assess restrictions in function that result from foot and ankle injury (Martin et al., 1999). It was intended for use with patients with a variety of ability levels. To develop the initial items, a literature review was conducted, patient feedback was gathered, other outcomes instruments were reviewed, and expert opinions were gathered from the APTA's Foot and Ankle Special Interest Group (Martin et al., 1999). There were initially 77 items, but the APTA's Foot and Ankle Special Interest Group reduced those down to 59. The group rated each item on a scale of -2 (not important) to +2 (very important) and items which had an average score of +1 were kept.

The FADI-Total was then divided into two subscales; activities of daily living (ADL; 22 activity items, 4 pain items) and sports-related activities (Sport; 8 items). The FADI-Total, therefore, consists of 34 items. The items are scored on a 0 – 5 scale. For most of the items a score of 0 indicates "unable to do" and a score of 4 indicates "no difficulty." For the pain scale, however, 0 indicates "none" and a score of 4 indicates "unbearable." The total possible points for the FADI-ADL are 104; total possible points for the FADI-Sport are 32. Each total score is then converted into a percentage and 100% indicates full function (Eechaute, Vaes, Van Aerschot, Asman, & Duquet, 2007; Hale & Hertel, 2005; Martin et al., 1999; Schlitz et al., 2009). Specifically, the FADI activities including standing, walking, stairs, squatting, heel raise, running, landing cutting, jumping, starting and stopping quickly, pain, ADLs, personal care, heavy work, light work, and desired participation.

Although the FADI is one of the most commonly used ankle outcomes instruments in the literature, and seems to have merit for those with high level of ability based on its sports sub-scale, it is not without measurement issues. Primarily, the original document in which the FADI was introduced by (Martin et al., 1999) was a research abstract which was missing key information. For instance, it is unknown exactly how the original 77 items were reduced. It was reported that a patient focus group, consisting of about 20 participants, was presented with the items to examine their understandability, but the details of that process were not provided. Psychometric analyses were not reported in the abstract, although the authors noted that Item Response Theory should be used for final item selection. Additionally, they noted that reliability and validity testing should also be included in final item selection, but the details of these procedures were not presented in the abstract.

A unique feature of the FADI is the Sport sub-scale. This scale was designed to measure outcomes in the physically active population and consists of 8 items which inquire about the patient's ability to run, land, cut, jump, squat, and perform their desired activities. It would seem, then, that this instrument should be effective at measuring the health status of high-ability individuals. However, Schlitz et al. (2009), in an unpublished thesis, found that almost half (120/244) of the participants who completed the FADI-Total, which includes the sport sub-scale, achieved a maximum score. One of the 75 participants from the Injured group, 12 of the 20 from the Healthy/Not Active group, 77 of the 110 from the Healthy/Active group, and 30 of the 39 Elite group had a maximum score on the FADI-Total. In regards to the FADI-ADL, 125 participants achieved a maximum score, and 140 participants had a maximum score for the FADI-Sport (Schlitz et al., 2009).

Since the original abstract by Martin, there has been only one publication that has examined the psychometrics of the FADI. In 2005, Hale and Hertel investigated the reliability and sensitivity of the FADI in subjects with chronic ankle instability (CAI). They examined a total of 50 participants (19 healthy, 31 with CAI). A sub-group (n = 16) of the CAI group participated in a 4-week ankle rehabilitation program. All participants completed the FADI and FADI-Sport during weeks 1, 4, and 7 of the study and were blinded to the previous responses. The authors calculated intraclass correlation coefficients (ICCs) between weeks 1 and 2, and weeks 1 and 7. They found moderate to good reliability at 1-week and 6-week intervals. The authors calculated the sensitivity of the FADI and FADI-Sport to CAI using a 2 x 2 mixed-model analysis of variance. They found that there were significant improvements in both the FADI and FADI-sport scores in the rehabilitation group from baseline to follow-up (Hale & Hertel, 2005). However, there was greater improvement in the FADI-Sport scores, indicating that it may be more sensitive to change. Paired *t*-tests were used to monitor changes in the affected ankles of the rehabilitation group from week 2 to week 7. Pearson product moment correlations were analyzed to examine the FADI and FADI-Sport scores at baseline. The analyses indicated that the sensitivity and reliability of the FADI and FADI-Sport are high enough to detect differences between healthy participants and participants with CAI. Furthermore, the authors concluded that the instruments are sensitive enough to detect changes in function after undergoing a rehabilitation program in participants with CAI (Hale & Hertel, 2005).

Although this study confirmed that the FADI shows evidence of validity for use with individuals with CAI, it did not progress towards the validation of its use with the high-ability population. The participants in this study were recreationally active at best; therefore, it did not examine the ceiling effect of the instrument. Additionally, a parametric analysis (paired *t*-test) was used to examine non-parametric data (FADI score changes over time). This is statistically improper and the results of this analysis should be interpreted cautiously.

Schlitz et al. (2009) found a significant difference between the FADI-ADL and the FADI-Sport, indicating that the FADI-Sport was more difficult. However, the results still indicate that although the sport sub-scale was designed for the physically active population and is more difficult than the FADI-ADL, there is still a ceiling effect for the FADI-Total. The instrument is not difficult enough to assess the activity limitations of the high-ability population.

The results of the Hale and Hertel (2005) study have been commonly misinterpreted, which has led to misuse of the FADI. Although the FADI was validated for use with patients with CAI, it was not intended to and has not been validated to identify persons who do and do not have CAI. Unfortunately, the use of the FADI as an inclusion criterion for participants with CAI of FAI has become quite common (C. Brown, 2011; Hubbard, 2008; Knapp et al., 2011; McKeon et al., 2008; Sefton et al., 2008; Sefton, Yarar, Hicks-Little, Berry, & Cordova, 2011). It has also been improperly used as a mechanism to measure FAI (Hubbard, Kramer, Denegar, & Hertel, 2007b). These practices should be cautioned as this was not the purpose for which the instrument was designed.

The FADI is not a perfect instrument. The original publication is a research abstract which lacks a great deal of detail regarding the development of the instrument, the instrument lacks thorough parametric testing on populations of various abilities, and it has a ceiling effect even when the sport sub-scale is included (Schlitz et al., 2009). However, compared to the other ankle and foot-specific self-reported outcomes instruments, the FADI seems to have the most merit and has made the most progress towards assessing the functional status of high-ability people.

Foot and Ankle Ability Measure (FAAM).

Two additional instruments were created from the FADI including the Foot and Ankle Ability Index (FAAI) and the Foot and Ankle Ability Measure (FAAM). The FAAI has 33 items, whereas the FADI has 34 total items. The validity and reliability of the FAAI have not been established (Pugia et al., 2001). The FAAM was developed to measure the physical ability of people with a variety of foot, ankle, and lower leg dysfunctions (Martin, Irrgang, Burdett, Conti, & Van Swearingen, 2005). Particularly, it assesses activities such as standing, walking, stairs, squatting, heel raise, running, landing cutting, jumping, starting and stopping quickly, pain, ADLs, personal care, heavy work, light work, desired participation.

The initial list of items was compiled from a literature review, physical therapists' opinions, and interviews with individuals with foot, ankle, and leg musculoskeletal injury. The items were then sent to experts from the American Physical Therapy Association (APTA) Foot and Ankle Special Interest Group for initial item reduction. The experts rated each item on a scale of -2 (not important) to +2 (very important) and items which had an average score of 1 (important) were included in the initial instrument. The experts were also asked if there should be ADL and Sport sub-scales for the instrument (Martin et al., 2005).

Scores for the initial FAAM were gathered from 1027 patients with a foot, ankle, or lower leg musculoskeletal disorder. Eigenvalues and factor loading trends, obtained via factor analysis, were used to identify factors. Factors were removed if they did not fit a one-factor model (Martin et al., 2005). Model fit data was calculated and the item difficulty and discrimination parameters were used to create the item characteristic curves. Items that did not have five distinct characteristic curves with one peak that represented the full range of ability were removed from the instrument (Martin et al., 2005).

Evidence of validity was established by separating the participants into two groups; one group was expected to change (*n* = 164, currently receiving physical therapy), the other was expected to not change (*n* = 79, no physical therapy for greater than one year). The group expected to change was currently receiving therapy, whereas the group not expected to change had received therapy at least one year prior. Both groups completed the FAAM and SF-36 two times, with approximately four weeks between time one and time two. The question "Over the past 4 weeks, how would you rate the overall change in your physical ability?" was added during the time two data collection (Martin et al., 2005).

Evidence of internal structure was established using factor analysis. Internal consistency was determined using Chonbach's alpha coefficient. A 95% confidence interval was used to assess the accuracy of the measurement at a specific point in time. Score stability evidence was established using intra-class correlation coefficients, which were calculated from the initial, as well as the 4-week, FAAM responses for the group that was not expected to change. The test-retest reliability over the 4 weeks was calculated using the intraclass correlation coefficient test re-test reliability coefficient to provide the standard error of the measure (SEM). The SEM was used to calculate the minimal detectable change (Martin et al., 2005).

To establish evidence of responsiveness, three analyses were performed; a two-way ANOVA with repeated measures, Guyatt's responsiveness index (GRI), and calculation of ROC curve. Evidence of convergent and divergent validity were established by examining the relationships between the FAAM and SF-36 physical function and physical component summary scores by calculating Pearson correlation coefficients (convergent); and by examining the relationships between the FAAM and SF-36 mental health and mental health summary scores (divergent). Sample size estimation revealed that 220 participants were needed to find a significant difference (Martin et al., 2005).

The results of the analyses revealed 21 items in the FAAM-ADL subscale and eight items in the Sports subscale. The items are scored on a 0 to 5 scale. A score of 0 indicates "unable to do" and a score of 4 indicates "no difficulty." For the pain scale, however, 0 indicates "none" and a score of 4 indicates "unbearable." The total possible points for the ADL scale are 84; total possible points for the sport subscale are 32. Each total score is then converted into a percentage. A score of 100% indicates full function.

Evidence of validity was found for the group that was expected to change, as well as for the group expected to remain the same, and the items showed appropriate stability and responsiveness (Martin et al., 2005). However, the instrument was not validated for use with high-ability individuals. The participants in this study were people who had or were currently partaking in physical therapy and were not classified as "high ability." Therefore, there is still a need for an instrument which is validated for and can accurately assess the functional status and outcomes of people across the entire ability spectrum.

Interestingly, it is not mentioned in the article that the FAAM was created from the FADI. One sleep item and four pain items were removed from the FADI to create the FAAM (Carcia, Martin, & Drouin, 2008), making these instruments nearly identical.

Carcia et al. (2008) attempted to provide evidence of validity for the FAAM in people with chronic ankle instability. They examined 30 NCAA Division II athletes (15 with CAI, 15 healthy). Each participant completed the FAAM along with global rating of function scale and an overall rating of function. Independent *t* tests were used to examine the differences in the FAAM subscales between the healthy and CAI groups and between the participants who rated their ankle as "normal" and "abnormal" in the overall rating of function. A Kendall tau rank correlation coefficient was used to assess the relationship between the FAAM scores and global rating scores for the FAAM subscales (Carcia et al., 2008).

The results of the study showed evidence of construct validity for the FAAM in athletes with chronic ankle instability. The FAAM ADL subscale scores were higher in participants who reported their ankles as normal, as compared to those who rated their ankles as nearly normal or abnormal. A moderate relationship was found between the ADL subscale and ADL global rating of function when the entire sample was included, but was weak when only the CAI group was included. A strong relationship was found between the sports subscale and the sports global rating of function when the entire sample was included, but was moderate when only the CAI group was included. The authors concluded that the FAAM is appropriate to use with athletes who have CAI (Carcia et al., 2008).

Interestingly, only one participant in the CAI group rated their ankle as abnormal. The rest rated their ankles as nearly normal or normal. The authors attributed this to the instructions of the FAAM, which only asks about the level of function over the last week. If the athlete did not experience any dysfunction in the previous week, then perhaps that is why they rated their ankles as nearly normal or normal. Post hoc testing showed that the CAI participants who ranked their ankles as normal had FAAM scores that were lower than scores of the healthy participants. Due to these score discrepancies, the authors felt that the CAI participants' ankles were truly abnormal. Additionally, these participants qualified for the CAI group, therefore indicating that they did not feel that their ankles were functioning completely normally (Carcia et al., 2008). However, it can be argued that although the athletes did, indeed, qualify for the CAI group, perhaps their ankles truly did not limit their normal function on a daily basis.

Shortcomings of this study include a small sample size, which limits the ability to generalize the results. The athletes were not examined over time, which eliminated the ability to measure reliability and responsiveness of the FAAM. However, the authors suggested that because the FAAM and FADI sports subscales are identical, and the sensitivity of the FADI has already been reported, then the FAAM would also exhibit sensitivity in athletes with CAI (Carcia et al., 2008). Although the results show that the FAAM has the capability to differentiate between healthy participants and participants with CAI, the instrument still cannot be used to assess change in functional status of high ability participants because some participants received a score of 100% on the sports subscale, indicating a ceiling effect. Future studies are still needed to appropriately measure the functional outcomes in high ability individuals.

There are many instruments that have been used to assess outcomes following foot and ankle injury. Although each has shortcomings, the FADI seems to have merit for use with individuals with high levels of physical ability. Therefore, the FADI was further explored in this research endeavor.

Conclusion

In conclusion, there are inconsistencies in the literature regarding the definition, inclusion criteria, and use of the terms "functional ankle instability" and "chronic ankle instability," which eliminates the possibility of comparing research findings. Additionally, there is a need for a valid instrument which can assess the clinical outcomes of high ability patients that is specific to the foot and ankle. Fortunately, new items can be added to existing instruments, rather than creating an entirely new instrument, to make the instrument more appropriate for individuals with high levels of physical ability. New items can be constructed by consulting a panel of experts and a group of individuals who fit the intended population of the instrument. However, before ankle outcomes following injury can be properly assessed, a model of ankle healthy must first be established. Therefore, the overall purpose of this dissertation was to develop a model for ankle function which allows researchers and clinicians to better assess patient outcomes following ankle injury. To accomplish this, two studies were conducted in order to gain insights from both content experts and elite athletes who have suffered foot and ankle injury.

CHAPTER III

STUDY 1

Although there is a need for instruments which accurately assess outcomes following ankle injury, there is not a valid ankle instrument for those with high levels of physical ability. Therefore, the collective purpose of this research project was to develop a model for ankle function which allows researchers and clinicians to better assess patient outcomes following ankle injury. To accomplish this, two studies were conducted utilizing two different designs. The first was conducted in order to gain insights from content experts, and the second to gain feedback from elite athletes who have suffered foot and ankle injury. This chapter provides a description of the methods, results, and discussion for Study 1.

Research Design

For Study 1, an observational, non-experimental Delphi design was used. The Delphi method is technique in which expert opinion is gathered to reach a consensus on a topic is. The Delphi method was initially used by the RAND Corporation during the Cold War to estimate the number of A-bombs necessary to increase the US military capability. The purpose of the Delphi method is to reach expert consensus without convening of the experts. The experts are questioned individually in a repeated fashion by means of interview or questionnaire until consensus is reached or until they begin to diverge from consensus. Convening of the experts is avoided in order to eliminate the influence that some experts may have on others. It promotes the expression of individual thoughts and ideas (Dalkey & Helmer, 1963). All procedures for this study were carried out in accordance to the University of Northern Iowa's Institutional Review Board policy. Detailed recruitment procedures and scripts, as well as consent documents, are included in the appendices (Appendix A1 – A3). No participants were excluded based on ethnic, racial, religious or cultural backgrounds.

Methods

The purpose of this study was to develop a model for a healthy ankle and address the inconsistencies of functional ankle instability.

Specific aim 1

To establish a definition of a healthy ankle based on feedback from a group of content experts.

Hypothesis 1

A consensus would be achieved as to the definition of a healthy ankle.

Specific aim 2

To determine the dysfunction and subsequent functional outcomes that result from ankle injury based on the interviews with the content experts.

Hypothesis 2

A consensus would be reached as to the dysfunction resulting from ankle injury.

Specific aim 3

To identify outcome instruments (tests, techniques, questionnaires) used to assess ankle dysfunction following injury based on the input from the content experts.

Hypothesis 3

A consensus would be reached regarding the outcome instruments that should be used to assess ankle dysfunction.

Specific aim 4

To establish a definition of functional ankle instability and to identify the resulting dysfunction.

Hypothesis 4

Although a consensus would be reached on the definition of FAI, no consensus will be reached regarding the resulting dysfunction.

Content Expert Recruitment

Potential content experts were identified based upon experience and expertise in the fields of ankle function, treatment, general outcomes assessment and research, in addition to a record of scholarship (more than five peer-reviewed publications in the content area), professional recommendations, and clinical experience (more than five years). There are no published guidelines as to what defines an expert, or the ideal number of experts that should participate (Steurer, 2011). There are general guidelines which state that the panel of experts should be heterogeneous in order to gather a variety of opinions (Keeney et al., 2001). Therefore, for the scope of this project, 16 experts were determined sufficient because it enabled a heterogeneous panel from a variety of fields. Once an expert was identified based upon the selection criteria, an email invitation (Appendix A1) was sent to request their participation on the expert panel. Experts that expressed interest in participating were sent an electronic consent form

(Appendix A2) and the interview questions. If the expert did not respond to the recruitment email, a follow-up telephone call was made within 48 hours following the initial email contact (Appendix A3). Recruitment ended once the panel included experts from the fields of athletic training, physical therapy, general medicine, and strength and conditioning; that had experience in ankle function, rehabilitation, FAI/CAI, outcomes research, had experience working with elite athletes; and were from both the clinical and academic settings.

Expert Demographics

Sixteen of 19 invited expert panelists participated in this study (Table 1). The experts were athletic trainers (8), physical therapists (2), dually credentialed athletic trainers/physical therapists (4), general medicine practitioners (1), and strength and conditioning specialists (1), who were currently practicing in the clinical (6), academic (8), or a combination of both clinical and academic (2) settings. The experts were selected based on their expertise in ankle function, treatment, general outcomes assessment, outcomes research, and clinical experience, and resided in the United States (14), Australia (1), and Ireland (1).

Table 1

Expert Panel's Demographic Information

Item	Total
Credentials	
Athletic Training (AT)	8
Physical Therapy (PT)	2
AT/PT (Dual credential)	4
General Medicine (MD)	1
Strength and Conditioning (CSCS, USA Weightlifting)	1
Highest Level of Education	
Bachelors	1
Masters	3
Doctorate	12
Years of Experience	
1 – 10	2
11 – 20	8
21 – 30	3
30+	3

Instruments

The initial round (Round 1) of data was collected using a series of structured interview questions which were carried out via telephone interview (Appendix A4). Responses were collected from each expert in the following content areas:

1. Definition of a healthy ankle
2. Functional characteristics of someone with a healthy ankle
3. Functional characteristics of someone with an unhealthy/acutely injured ankle

4. Assessment of functional characteristics
5. Definition of FAI
6. Functional characteristics of someone with FAI
7. Assessment of functional characteristics of someone with FAI
8. Appropriateness of the term FAI
9. Definition of CAI
10. Functional characteristics of someone with CAI
11. Assessment of functional characteristics of someone with CAI
12. Appropriateness of the term CAI

These content areas were selected to gather information regarding the definition, characteristics, and assessment techniques of a healthy ankle, unhealthy ankle, FAI, and CAI. Subsequent rounds of data collection were carried out via Survey Monkey™ (www.surveymonkey.com, 1999-2011).

Procedures

The Delphi method was used in Round 1 to gather expert opinion regarding ankle function. The Delphi method involves interviews with content experts in effort to reach a consensus regarding the topic of interest. National and international experts were invited to participate in the interviews. The interview questions were read aloud over the phone to the content experts. They also received the questions electronically prior to the phone interview. Once the responses from the telephone interview questions were analyzed and reduced, all subsequent rounds of data collection were administered via Survey Monkey. Because the data was collected in rounds, and the content of subsequent rounds were based upon each preceding round, the specific procedures and analyses will be presented in a step-by-step manner.

Round 1

The telephone interviews were recorded using an Olympus WS-210S Digital Voice Recorder (3500 Corporate Parkway, Center Valley, PA) and a telephone call recording device (VEC, TRX-20 D). The total interview time ranged from nine minutes, 30 seconds, to 29 minutes, three seconds. The average interview time was 18 minutes, six seconds. All sixteen experts participated in Round 1.

Expert orientation and training. To ensure that the experts understood the purpose of the study, their role in the study, and the procedures of the interview, the experts were oriented at the beginning of the telephone interview (Appendix A4). They were asked to state their opinion about their understanding of ankle function. After each response, the experts were asked if their answer was complete before proceeding to the next question. Following the orientation, the experts were provided the opportunity to ask questions regarding the study.

Data analysis. The telephone interviews were transcribed verbatim and analyzed descriptively. The responses were compiled and reduced qualitatively to detect common themes and to categorize the data; and frequencies were calculated quantitatively. To minimize researcher bias, and to maximize trustworthiness of the data, two other members of the research team who were experienced researchers assisted in the data reduction process and reviewed the final reductions. All responses provided by the experts in Round 1 appeared in Round 2 as response choices to the interview questions. For instance, all of the reduced responses to question one in Round 1 became the response choices for question one in Round 2.

Round 2

Procedures. The reduced data from Round 1 were entered into an electronic survey (Survey Monkey, Round 2). The only set of responses that did not appear in Round 2 were the responses to the questions regarding assessment techniques because consensus on the functional characteristics and/or limitations had not yet been reached. An example question is: "Please indicate if you agree or do not agree that each of the following items should be included in the functional assessment of acute ankle injuries." The experts were then asked to respond to each item under each question to indicate if they "agree" or "disagree" that the item should be included as a component of the question or as part of the definition. Example items include: "plantarflexion range of motion, anterior drawer test, Star Excursion Balance Test, etc." All sixteen experts participated in Round 2.

Data analysis. Although there is no universally accepted criterion for "consensus," 80% agreement is often considered acceptable (Boulkedid et al., 2011) and has been used in other Delphi studies (Chung et al., 2012) and thus, was used in this study. From Round 2, and each subsequent round, the items that reached 80% or greater consensus indicated that it should be included (agree) were retained and were not posed to the experts again for further consideration, but appeared in the final definition or as a component of the construct in question. Conversely, items that reached 80% or greater consensus indicated that it should not be included (disagree) were discarded and did not appear in subsequent rounds. Items that did not reach 80% or greater consensus to be either included or not included were also eliminated because they did not achieve consensus.

Round 3

Procedures. In Round 3, the panel was asked to "agree" or "disagree" with the definition of a healthy ankle that was established based on the experts' feedback in Round 2. Data from Round 2 regarding *only* a healthy ankle were re-presented to the panel in this round. Experts were also presented with each of the items that the panel agreed should be included in the definition of a healthy ankle. For the remainder of the study, the experts were provided with a greater number of short surveys with fewer questions, rather than a smaller number of long surveys with several questions in order to prevent attrition. All 16 experts participated in Round 3.

Data analysis. Round 3 data were analyzed descriptively as percentages of agreement and disagreement with the proposed definition of a healthy ankle.

Round 4

Procedures. In Round 4, the panel was asked to "agree" or "disagree" with each of the assessment techniques that were identified in Round 1. Additionally, the panel was presented with an update of Round 3 results, the functional characteristics of a healthy ankle that were agreed upon in Round 2, as well as the functional characteristics of acutely injured/unhealthy ankles that were also agreed upon in Round 2. Fifteen of the 16 experts participated in Round 4.

Data analysis. Round 4 data were analyzed descriptively as percentages of agreement and disagreement with each of the assessment techniques.

Round 5

Procedures. From the feedback gathered in Round 2, definitions of functional ankle instability and chronic ankle instability (CAI) were created. In Round 5, the panel was asked to "agree" or "disagree" with the definitions of FAI and CAI. Within this round, the panel was also provided an update on the feedback regarding the functional characteristics or limitations of someone with FAI, different titles or names used for FAI, a name more appropriate than FAI, if FAI and CAI are thought to be the same, the functional characteristics or limitations of someone with CAI, different titles or names used for CAI, and if "persistent ankle instability" is an appropriate substitute. Fifteen of the 16 experts participated in Round 5.

Data analysis. Round 5 data were analyzed descriptively as percentages of agreement and disagreement with each of the definitions of FAI and CAI.

Results and Discussion

The purpose of this research project was to develop a model for ankle function which allows researchers and clinicians to better assess patient outcomes following ankle injury. Participant demographics, results, and discussion are presented in this section. The results will be presented with the original question to which they pertain. In addition to responding to the questions in each round of surveys (Rounds 2 – 5), the experts were asked to provide written comments at the end of each survey. Due to the lengthiness of the comments, they are included as an appendix (Appendix C).

Definition of a Healthy Ankle

In Round 1, the experts were asked: "How do you define a healthy/normal/non-injured ankle?" Table 2 represents the results of Round 1 interviews, as well as Round 2 survey questions (agree/disagree) because the reduced responses from Round 1 were provided back to the experts in Round 2. Experts provided various responses and most experts listed components of a healthy ankle, rather than provide a singular definition. The list of items that were agreed upon in Round 2 were then compiled and reduced and a definition of a healthy ankle was created. After reducing redundant items from Round 2, 16 items remained to be included in the definition of a healthy ankle: (1) normal gait, (2) normal coordination of movement, (3) no pain, (4) no complaints, (5) no perceived functional deficits, (6) no restrictions in desired participation, (7) no compensations in movement activities related to the ankle, (8) normal range of motion, (9) normal strength, (10) normal arthokinematics and osteokinematics, (11) normal circulation, (12) normal sensation, (13) no swelling, (14) no synovial changes, (15) no pathology, and (16) no osteoarthritis. Because it was not practical to incorporate all 16 items into a definition, the 16 items were fit into five broader inclusive categories including functions, participation, impairments, pathology, and pain. Although pain is generally considered an impairment, the research committee decided that pain was important and unique enough to be represented as an individual component of the definition. Therefore, the data indicated that the definition of a healthy ankle is: "A person with a healthy ankle(s), either through self-report or clinical measurement, presents with full functional capacity and participation status, without pathology, pain, or impairments relative to the ankle(s)."

This definition was then presented back to the expert panel in Round 3. Eighty-seven percent (14) of the panel agreed with this definition of a healthy ankle and, therefore, data collection was concluded for the definition.

However, multiple experts emphasized that some components of a healthy ankle are not limited to only ankle function. For instance, one could have a healthy ankle, but their functional capacity could be limited by another injury. Another expert commented that they feel there is a difference between healthy/normal and non-injured; that an abnormality could be present in the absence of injury. Additionally, some experts noted that pain is an impairment and suggested that pain not be listed as a separate component of the definition. These comments will be further discussed later in this chapter.

Functional Characteristics of a Healthy/Normal/Non-Injured Ankle

The second interview item was: "Describe the functional characteristics of someone with a healthy/normal/non-injured ankle." The responses are listed in Table 3. These responses were then presented back to the panel in Round 2 using an electronic survey and the experts were asked to agree or disagree with each response (Table 3). The functional characteristics that reached 80% consensus were retained in the final list. Those that did not reach 80% were eliminated. The final list of functional characteristics of a healthy ankle included: (1) normal gait pattern, (2) normal balance, (3) normal neuromuscular control, (4) no disability, (5) no apprehension or fear of performing desired activities, (6) no patient reports of giving way, (7) patient can ambulate stairs, (8) patient can run, (9) patient can perform all desired activities (including sport, work, and

ADLs), (10) patient feels they are at pre-injury levels, and (11) it depends on what they want to do (it is based on patient's perspectives and goals).

Table 2

"How do you define a healthy/normal/non-injured ankle?"

Initial Responses	Round 2 Responses	
	Agree (Total)	Disagree (Total)
Normal range of motion	16*	0
Normal muscle strength	15*	1
Normal gait pattern	14*	2
Normal joint mechanics	15*	1
Normal arthrokinematics	16*	0
Normal joint position	13*	3
Normal joint stability	16*	0
Normal circulation	14*	2
Normal sensation	15*	1
Normal coordination of movement	14*	2
No pain	15*	1
No complaints	15*	1
No complaints of instability	16*	0
No perceived functional deficits	16*	0
No restrictions in desired participation	15*	1
No history of ankle injury	7	9
No mechanical laxity	14*	2
No pathology	16*	0
No swelling	15*	1
No synovial changes	13*	3
No osteoarthritis	16*	0
No compensations in movement activities related to the ankle	13*	3
Patient can complete full functional assessment	13*	3
Patient can perform deep body squats without problems	10	6

* = 80% consensus was reached

Table 3

"Describe the functional characteristics of someone with a healthy/normal/non-injured ankle."

Initial Responses	Round 2 Responses	
	Agree (Total)	Disagree (Total)
Normal active range of motion	12	4
Normal passive range of motion	12	4
Normal resistive range of motion	12	4
Full dorsiflexion	9	7
Normal inversion and eversion	10	6
Normal gait pattern	15*	1
Normal muscular endurance	12	4
Normal muscle strength	12	4
Normal balance	14*	2
Normal proprioception	11	5
Normal neuromuscular control	14*	2
Normal arthrokinematic motion	12	4
Normal perceived postural stability	12	4
No disability	14*	2
No apprehension or fear of performing desired activities	14*	2
No swelling	12	4
No pain	11	5
No patient reports of giving way	15*	1
Patient can bear weight 8x's body weight, age dependent	4	12
Patient can get in to squat with heel on ground	7	9
Patient can ambulate stairs	15*	1
Patient can run	14*	2
Patient can perform all desired activities (including sport, work, and ADLs)	13*	3
Patient feels they are at pre-injury levels	14*	2
Patient feels stable	12	4
Ankle is stable	12	4
It depends on what they want to do (it is based on patient's perspectives and goals)	14*	2

* = 80% consensus was reached

Many comments were provided regarding this question. For instance, as pointed out by one expert, the impairments may be required for function, even though the impairments themselves are not functions. For example, "patients can be functional without full range of motion," or "they can have pain and still be functional," as also pointed out by one of the experts in their comments. Another expert mentioned that an appropriate addition to each of the characteristics might be "as related to the ankle" to clearly identify the functional ability of the ankle. Other comments and issues that arose pertain to the term "decreased." As in the list of characteristics of a healthy ankle, the reference point for "normal" was not clearly identified. Also, it is not clear if a patient must have all or just a single characteristic on the list to be considered "unhealthy." Another comment pointed out that, items such as "decreased confidence" may be due to factors outside of the ankle, so an addition may be to specify "as related to the ankle" to each characteristic. These comments will be further discussed later in this chapter.

Functional Characteristics/Limitations of an Unhealthy/Acutely Injured Ankle

The third interview item was: "Describe the functional characteristics or limitations of someone with an unhealthy/acutely injured ankle." Responses were then presented back to the panel in Round 2, and the expert panel's responses and results can be found in Table 4. Only the characteristics that reached an 80% agreement were retained, those that did not were eliminated. The final list of functional characteristics of an unhealthy/acutely injured ankle included the following: (1) pain, (2) loss of range of motion, (3) loss of dorsiflexion, (4) loss of dorsiflexion/plantarflexion, (5) unable to engage in desired level of function, (6) unable to cut, (7) unable to do back pedals, (8) unable to jump, (9) unable to ascend and descend stairs, (10) unable to change directions, (11) unable to perform transitions, (12) unable to squat, (13) unable to run, (14) unable to plant, (15) unable to land from a jump, (16) unable to engage unpredictable surfaces, (17) unable to perform ADLs, (18) unable to single leg balance, (19) decreased mobility, (20) abnormal joint mechanics, (21) swelling, (22) neuromuscular inhibition, (23) altered neuromuscular drive, (24) reflex inhibition, (25) decreased proprioception, (26) decreased coordination, (27) decreased balance, (28) decreased postural stability, (29) movement compensation, (30) apprehension, (31) decreased confidence, and (32) it depends on patient's desired or normal level of activity. Comments that emerged from this question were similar to that of the functional characteristics of a healthy ankle and will be discussed later in this chapter.

Table 4

"Describe the functional characteristics or limitations of someone with an unhealthy/acutely injured ankle."

Initial Responses	Round 2 Responses	
	Agree (Total)	Disagree (Total)
Pain	13*	3
Loss of range of motion	13*	3
Loss of dorsiflexion	13*	3
Loss of dorsiflexion/plantarflexion	13*	3
Loss of inversion/eversion/pronation/supination during gait	12	4
Unable to engage in desired level of function	16*	0
Unable to cut	15*	1
Unable to do eccentric exercise	12	4
Unable to do slow-downs	12	4
Unable to do back pedals	14*	2

(Table continues)

Initial Responses	Round 2 Responses	
	Agree (Total)	Disagree (Total)
Unable to jump	16*	0
Unable to ascend and descend stairs	15*	1
Unable to change directions	16*	0
Unable to perform transitions	15*	1
Unable to squat	14*	2
Unable to run	15*	1
Unable to plant	14*	2
Unable to land from a jump	15*	1
Unable to engage unpredictable surfaces	15*	1
Unable to perform ADLs	14*	2
Unable to single leg balance	14*	2
Decreased mobility	14*	2
Decreased muscle strength	12	4
Decreased muscular endurance	11	5
Limited power	12	4
Muscles have increased tone	7	9
Increased range of motion	8	8
Increased pronation/supination	8	8
Increased joint laxity	12	4
Abnormal joint mechanics	13*	3
Capsular tightness	9	7
Swelling	13*	3
Neuromuscular inhibition	14*	2
Altered neuromuscular drive	13*	3
Reflex inhibition	13*	3
Decreased proprioception	14*	2
Decreased coordination	15*	1
Decreased balance	16*	0
Decreased postural stability	15*	1
Movement compensation	14*	2
Apprehension	13*	3
Decreased confidence	13*	3
Depends on patient's desired or normal level of activity	15*	1
Depends on the severity of injury	11	5
Costs related to income/health care	4	12

* = 80% consensus was reached

Assessment Techniques

In the interview, the experts were asked: "How do you assess functional characteristics of the ankle (for injured or healthy ankles)?" The responses were compiled and returned to the experts in Round 4. These results are listed in Table 5. Because only 15 of the 16 experts responded to Round 4, 80% consensus was achieved at 12 responses. After final reduction, the ankle assessment techniques included: (1) gait analysis, (2) weight-bearing ability/willingness, (3) ability to walk on uneven ground, (4) ability to jog without pain, (5) running, (6) stair ambulation, (7) watch patient during activity, (8) jumping, (9) single leg jumping, (10) hopping patterns, (11) agility tests, (12) ability to cut without pain, (13) changing directions, (14) Figure-8 test, (15) single leg squatting, (16) double leg squatting, (17) balance assessment, (18) single leg stance, (19) single leg stance for time (eyes closed), (20) Star Excursion Balance Test, (21) ability to land evenly from a jump, (22) jump landing with stabilization holds, (23) single leg jump landings, (24) range of motion assessment, (25) active range of motion assessment, (26) passive range of motion assessment, (27) dorsiflexion range of motion assessment, (28) special tests, (29) ligamentous testing of the ankle, (30) anterior drawer test, (31) talar tilt test, (32) proximal and distal tib/fib joint assessment, and (33) history/subjective report. One addition to this list suggested by an expert in his comments was to add "can perform without pain" to the assessments. One expert commented: "What I really like about this is that it seems as though laboratory-oriented, clinician-oriented, patient-oriented measures have emerged as being relevant for assessment. This is a great sign that we're looking for context among the patient, clinician, and researcher!"

Table 5

"How do you assess functional characteristics of the ankle (for injured or healthy ankles)?"

Initial Responses	Round 4 Responses	
	Agree (Total)	Disagree (Total)
Motion analysis	9	6
Gait analysis	14*	1
Weight-bearing ability/willingness	15*	0
Ability to walk on uneven ground	13*	2
Ability to jog without pain	12*	3
Running	15*	0
Stair ambulation	14*	1
Functional analysis	10	5
Functional tests on different surfaces	11	3
Watch patient during activity	12*	3
Willingness to do activities	10	5
Jumping	15*	0
Single leg jumping	15*	0
Ability to single leg jump for distance	10	5
Double leg jumping	10	5
Standing jumps	10	5
Hopping patterns	14*	1
Hop for distance	11	4
3 Hop Test	10	5
Hop and Stop Test	8	7
Ability to back pedal without pain	10	5
Lunge tests for dorsiflexion ROM	11	4
Agility tests	14*	1
Ability to cut without pain	12*	3
Changing directions	15*	0
Obstacle course	7	8
Figure-8 Test	14*	1
Ability to repeat activity overtime without breaking down	10	5
Single leg squatting	14*	1
Double leg squatting	12*	3
Deep squatting with full dorsiflexion	9	5
Postural control	10	5
Balance assessment	15*	0
Single leg stance	14*	1

(Table continues)

Initial Responses	Round 4 Responses	
	Agree (Total)	Disagree (Total)
Single leg stance for time (eyes closed)	13*	2
Tandem stance	8	7
Tandem stance for time (eyes closed)	9	6
Rhomberg Test	4	11
Balance Error Scoring System	11	4
Star Excursion Balance Test	14*	1
Ability to land evenly from a jump	13*	2
Jump landing with stabilization holds	12*	3
Single leg jump landings	12*	3
Double leg jump landings	9	6
Landing Error Scoring System	8	7
Step-to-stabilize	9	6
Manual Muscle Tests	11	4
Peroneal strength test	11	4
Gluteus maximus strength test	5	10
Gluteus medius strength test	5	10
Tibialis anterior strength test	8	7
Gastrocnemius strength test	10	5
Soleus strength test	8	7
Tibialis posterior strength test	8	7
Strength tests using a hand-held dynamometer	3	12
Isokinetic tests	3	12
Range of motion assessment	13*	2
Active range of motion assessment	13*	2
Passive range of motion assessment	12*	3
Dorsiflexion range of motion assessment	13*	2
Special tests	13*	2
Ligamentous testing of the ankle	12*	3
Anterior Drawer Test	13*	2
Kleiger's (External Rotation)Test	10	5
Talar Tilt Test	13*	2
Squeeze Test	8	7
Tap Test	7	8
Arthrometer testing	6	9
Modified Ottawa Ankle Rules	10	5
Arthrokinematic assessment	11	4
Talocrural joint assessment (Maitland)	11	4
Proximal and distal tib/fib joint assessment	12*	3

(Table continues)

Initial Responses	Round 4 Responses	
	Agree (Total)	Disagree (Total)
Mulligan fibular AP excursion of inferior tib/fib joint	10	5
Neurological exam	9	6
Circulation assessment	7	8
Sensation assessment	9	6
Swelling measurement	8	7
EMG	3	12
Footwear assessment	10	5
Palpate before and after activity for tenderness	5	10
Ask the coach if there has been a change in performance	6	9
History/subjective report	14*	1
Foot and Ankle Disability Index	8	7
Foot and Ankle Ability Measure	10	5
Cumberland Ankle Instability Tool	10	5
Foot and Ankle Outcome Scale	7	8
SF-36	8	7

* = 80% consensus was reached

Definition of Functional Ankle Instability

In the initial Round 1 interview, the experts were asked: "What is your definition of functional ankle instability?" Similar to their response to the definition of a healthy ankle, they tended to provide lists of components rather than a singular definition. The experts' responses to the question from Round 1, as well as the subsequent results from Round 2 are presented in Table 6.

Based upon the experts' responses to the Round 2 questioning of the definition of functional ankle instability, the working definition of functional ankle instability was established as "Functional ankle instability is a recurrent sense of giving way of the ankle." This definition was presented to the experts in Round 5 and 12 (80%) of the 15 respondents in this round agreed with the definition of FAI. One expert opted out of this round of questioning. Additionally, there were very strong comments provided in support of FAI, in contrast to CAI, and vice versa. For instance, in one comment an expert stated in regards to FAI: "....if you must use that term." In another comment a different expert stated CAI: "....is a bad term...." and FAI "....is the better term...." Another expert commented that they would have liked to have been able to select "not necessarily." These comments will be further discussed later in this chapter.

Table 6

"What is your definition of functional ankle instability?"

Initial Responses	Round 2 Responses	
	Agree (Total)	Disagree (Total)
I wish I knew	4	12
I am not sure what functional ankle instability is	3	13
I don't like the term "functional ankle instability"	6	10
Pain	3	13
Inability to perform upper level activities without pain	5	11
Recurrent ankle sprains	12	4
History of at least one ankle sprain	1	15
History of at least one ankle sprain in the previous year	1	15
History of a moderate to severe ankle sprain at least 12 months in the past that required crutches or NWB for at least 3 days	2	14
Subjective complaints	6	10
Subjective complaints that can be reproduced clinically	6	10
Recurrent sense of giving way	14*	2
At least two episodes of giving way	8	8
Symptoms need to occur after perceived recovery	6	10
Perceived instability	12	4

(Table continues)

Initial Responses	Round 2 Responses	
	Agree (Total)	Disagree (Total)
Inability to participate in desired activities	10	6
Subjective feeling that ankle prevents them from participating in something	12	4
Fear avoidance/kinesiophobia	8	8
Lack of confidence	7	9
Loss of range of motion	4	12
Abnormal gait pattern	4	12
Swelling	4	12
Ligament laxity	4	12
Excessive joint motion	4	12
Decreased muscle strength	4	12
Decreased peroneal muscle strength	3	13
Decreased proprioception	4	12
Neuromuscular re-education issues	6	10
Sensorimotor deficits	10	6
FADI score of less than or equal to 90	4	12
FADI sport subscale score of 80% or lower	5	11
CAIT score of less than or equal to 24	7	9

* = 80% consensus was reached

Functional Characteristics of Functional Ankle Instability

During the interview, the experts were asked to: "Describe the functional characteristics or limitations of someone with functional ankle instability; what are the functional costs or consequences?" Their initial responses, and the results of this question in Round 2 (agree/disagree) are presented in Table 7. The functional characteristics that reached 80% consensus were retained. Therefore, the list of functional characteristics of FAI included: (1) unable to plant/cut/change directions, (2) unable to hop, (3) unable to function at desired level of activity, (4) unable to participate in desired activities, (5) unable to complete sport-specific activities, (6) difficulty walking on uneven ground, (7) difficulty with single leg landing, (8) decreased confidence, (9) repetitive episodes of giving way, (10) inhibited neuromuscular control, (11) diminished performance, (12) there is a continuum, and (13) it depends on what they want to do. The comments provided in response to this question were similar to the comments provided in regards to the definition of FAI and will be discussed later in this chapter.

Table 7

"Describe the functional characteristics or limitations of someone with functional ankle instability; what are the functional costs or consequences?"

Initial Responses	Round 2 Responses	
	Agree (Total)	Disagree (Total)
There are no limitations	1	15
It is the same as for acute ankle injury	3	13
Pain	9	7
Decreased muscular endurance	9	7
Unable to plant/cut/change directions	14*	2
Unable to jump or hop	13*	3
Unable to function at desired level of activity	16*	0
Unable to bear weight	6	10
Unable to participate in desired activities	14*	2
Unable to complete sport-specific activities	13*	3
Difficulty ambulating stairs or ladders	11	5
Difficulty walking on uneven ground	13*	3
Difficulty with single leg landing	13*	3
Inhibited eccentric muscle contraction	8	8
Repetitive ankle sprains	10	6
Repetitive episodes of giving way	14*	2
Poor ligament stability	7	9
Abnormal gait	11	5
Fear avoidance/apprehension	12	4
Decreased confidence	13*	3
Limited Star Excursion Balance Test	12	4
Limited Balance Error Scoring System	11	5
Inhibited balance (single and double leg)	12	4

(Table continues)

Initial Responses	Round 2 Responses	
	Agree (Total)	Disagree (Total)
Inhibited postural control	10	6
Inhibited neuromuscular control	13*	3
Inhibited proprioception	11	5
Inhibited muscular endurance	10	6
Perceived weakness	10	6
Diminished performance	16*	0
Decreased range of motion	10	6
Range of motion is shifted away from dorsiflexion and towards plantarflexion	9	7
Increased inversion	7	9
Chronic swelling	7	9
Decreased flexibility	7	9
Synovial changes	7	9
Arthrokinematic changes	10	6
Ligament laxity	9	7
Degenerative changes	8	8
There is a continuum	13*	3
It depends on what they want to do	14*	2

* = 80% consensus was reached

Different Title or Name for Functional Ankle Instability

The experts were asked in the initial interview: "Do you have a different title or name that you use for functional ankle instability?" The responses to that question are presented in Table 8, along with the responses to the same question in Round 2. Based upon the lack of consensus, there are no other terms that are used in place of FAI.

Name More Appropriate Than Functional Ankle Instability

During the Round 1 interview, the experts were asked: "Do you feel that there is a name more appropriate than functional ankle instability?" The responses to this question, and the responses to the same question in Round 2, can be viewed in Table 9. Based upon the experts' response, there is not a name more appropriate than FAI.

Table 8

"Do you have a different title or name that you use for functional ankle instability?"

Initial Responses	Round 2 Responses	
	Agree (Total)	Disagree (Total)
No, I use the term functional ankle instability	9	7
Not sure what term I use	3	13
Unstable ankle	7	9
Chronically unstable ankle	5	11
Chronic ankle instability	11	5
I use chronic ankle instability, but feel it means the same thing as functional ankle instability	5	11
Injured ankle	1	15
Depends on my audience	6	10

* = 80% consensus was reached

Are FAI and CAI the Same Thing?

The experts were asked: "Do you feel that functional and chronic ankle instability are the same thing?" in the first round of interviews, consensus was not reached on a singular response, but the majority of the experts (9, 56.2%) agreed that CAI is the umbrella term and FAI is a component. Additionally, three experts responded "No," for a total of 13 (81.3%) experts who agreed that CAI and FAI are not the same. The remaining four experts (25%) responded "yes" indicating that they agree that CAI and FAI are the same.

Table 9

"Do you feel that there is a name more appropriate than functional ankle instability?"

Initial Responses	Round 2 Responses	
	Agree (Total)	Disagree (Total)
I feel functional ankle instability is appropriate	12	4
I am not sure	3	13
I wish I could think of one	1	15
Unstable ankle	4	12
Chronic ankle sprainer	3	13
Chronic ankle dysfunction	7	9
Chronic ankle instability	8	8
Recurrent ankle dysfunction	7	9
Recurrent ankle instability	5	11
Ankle dysfunction	5	11
Somatic dysfunction	1	15
Perceived ankle instability	5	11
Functional instability of the ankle joint	4	12

* = 80% consensus was reached

Definition of Chronic Ankle Instability

In the Round 1 interview, the experts were asked: "What is your definition of chronic ankle instability?" Their responses, as well as the responses to the same question in Round 2 (agree/disagree) are presented in Table 10. Similar to the previous definition questions, the experts provided lists of components, rather than a single definition. From the Round 2 responses, the working definition of CAI was established as "Chronic ankle instability is recurrent episodes of ankle instability and sprains." This definition was then given back to the experts in Round 5 for their feedback and 11 (68.8%) of the 16 respondents agreed with the definition, but did not reach the 80% standard for consensus to be established. The comments that the experts provided in regards to this question were similar to the comments provided for the definition of FAI and will be discussed later in this chapter.

Table 10

"What is your definition of chronic ankle instability?"

Initial Responses	Round 2 Responses	
	Agree (Total)	Disagree (Total)
It is not different than FAI	3	13
Series of injuries to the ankle over time	10	6
It is a syndrome	8	8
Recurrent episodes of instability	15*	1
Perceived instability	10	6
Recurrent sprains	14*	2
First sprain occurred greater than one year ago and there are still symptoms related to that injury	5	11
Pain	8	8
Swelling	8	8
Apprehension	9	7
Functional deficits	12	4
Self-reported functional deficits	11	5
Functional ankle instability that occurs for a long period of time	9	7
Ligament laxity	7	9

* = 80% consensus was reached

Functional Characteristics of Chronic Ankle Instability

Another question that the experts were asked in the Round 1 interview was: "What are the functional characteristics of someone with chronic ankle instability; what are the functional costs or consequences?" The responses to the initial question, along with the results from the same question in Round 2 (agree/disagree) are in Table 11.

There were no functional characteristics that reached consensus for inclusion for CAI, but consensus was reached that CAI is different than FAI.

Table 11

"What are the functional characteristics of someone with chronic ankle instability; what are the functional costs or consequences?"

Initial Responses	Round 2 Responses	
	Agree (Total)	Disagree (Total)
I am not sure if there is a performance decrease or cost	3	13
Same as functional ankle instability	12	4
Same as functional ankle instability, just a longer period of time	3	13
It depends on the sport/activity they participate in	11	5
Everyone is different	11	5
Ankle ligament laxity	5	11

* = 80% consensus was reached

Different Name for Chronic Ankle Instability

During the Round 1 interview, the experts were asked: "Do you have a different title or name that you use to describe chronic instability?" The responses were then sent back to the experts for further feedback in Round 2. Both Round 1 and Round 2 (agree/disagree) responses are listed in Table 12. The majority of the experts (13, 81.3%) use the term CAI.

Is "Persistent Ankle Instability" an Appropriate Substitute?

The last interview question in Round 1 was: "Do you feel that "persistent ankle instability" is an appropriate substitute for the title or name that you use (e.g. functional ankle instability, chronic ankle instability, ankle instability, etc.)?" There was no response that achieved 80% consensus in regards to this question, but the majority of the experts (9, 56.2%) agreed that "persistent ankle instability" is not an appropriate substitute. Additionally, four experts (25%) indicated that they agree that "persistent ankle instability" is an appropriate substitute, while three (18.8%) of the experts indicated that "it all means the same thing, just different words".

Table 12

"Do you have a different title or name that you use to describe chronic instability?"

Initial Responses	Round 2 Responses	
	Agree (Total)	Disagree (Total)
No, I use the term chronic ankle instability	13	3
Unstable ankle	3	13
Chronic ankle sprainer	2	14
Instability	3	13
Ankle instability	8	8
Insufficiency	2	14
Impairment	2	14
Chronic functional ankle instability	3	13
Recurrent ankle dysfunction	6	10
Recurrent ankle instability	6	10
Ankle injury	1	15
Depends on my audience	5	11

* = 80% consensus was reached

Discussion

Definition of a Healthy Ankle

In order to properly measure health care outcomes, the element of health that is of interest must be determined, along with an instrument to measure it. This can be one of the most challenging steps in the process of assessing outcomes. One element of health that is often addressed is level of health, which has been described as function (Nagi, 1965), activity limitation (World Health Organization, 2002), disability (Nagi, 1965), impairment (Nagi, 1965; World Health Organization, 2002), and participation status (World Health Organization, 2002). Models such as these have been created to provide a common language that can be used by all health care practitioners. The most important component of these models is that they focus broadly on the patient's health status, including social well-being, rather than simply focusing on the physical limitations of the disability (Jette, 2006; Snyder et al., 2008).

The most recent model of health, established by the World Health Organization International Classification of Function (World Health Organization, 2002), includes assessment of impairments, which include clinician-measured outcomes, such as range of motion, strength, and participation. Participation refers to the ability of the patient to function within their environment, including their social environment (World Health Organization, 2002). These components were incorporated into the definition of a healthy ankle. This was an important step because before outcomes can be accurately measured, health must first be defined.

The definition of a healthy ankle had not been previously established; rather the literature focused on determination and measurement of ankle injury and instability. Therefore, the specific aim of this study in regards to the definition of a healthy ankle was to establish a definition of a healthy ankle and it was hypothesized that a consensus would be achieved. Although consensus was reached for the definition of a healthy ankle, issues remain. As multiple experts pointed out, some components are not limited to only ankle function. For instance, one could have a healthy ankle, but their functional capacity could be limited by another injury. Another expert commented that they feel there is a difference between healthy/normal and non-injured; that an abnormality could be present in the absence of injury. Additionally, some experts noted that pain is an impairment and suggested that pain not be listed as a separate component of the definition. However, pain remained in the definition because it is such an important component, and the disablement model terminology is not yet universal for all health care professionals to recognize that pain is an impairment. Although this definition may not be the final definition, it shows progress towards establishing a common definition of ankle health.

Similar to the specific aims of this study, a group of experts convened in Vienna, Australia in 2001 to review all components of concussion, including the definition (Aubry et al., 2002). Although a definition of concussions had already been established, it had limitations, so one of the objectives of this forum was to revise the definition. The expert panel was successful in their attempt to come to a consensus on multiple aspects of concussion, including the definition. Much like the current study sought the input of a panel of experts regarding the definition of ankle health, this study used a panel of experts as well, but they used an in-person forum, rather than telephone interviews and surveys. It is important to consider that a group of experts was tasked with determining the definition of a potentially life-threatening condition, which was successful. This supports the use of the Delphi method to establish a definition of a healthy ankle because it was simply not feasible to bring all of the experts together in a forum.

Functional Characteristics of a Healthy Ankle

The second specific aim of this study was to determine the dysfunction and subsequent functional outcomes that result from ankle injury. It was hypothesized that a consensus would be reached. Although the experts were asked to describe the functional characteristics of a healthy ankle, some listed items were not functions, but rather impairments (i.e., range of motion, strength, pain, balance, etc.). However, as pointed out by one expert, the impairments may be required for function, even though the impairments themselves are not functions. This is indicative of the fact that the disablement model language is not universal to all health care professionals. Whereas some experts identified that some items on the list were not functions, others perhaps did not know the difference between a function and an impairment. Some experts did not select the impairments for inclusion in the final list of characteristics and, therefore, many of these impairments are not included in the final list of characteristics of a healthy/normal/non-injured ankle. It is clear that we must continue to strive towards the use of a common language among all health care professionals and that the disablement framework, as established by the WHO ICF, must be disseminated throughout the health care profession.

According to the WHO, health is defined as "state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (http://www.who.int/about/definition/en/print.html. Accessed May 17, 2012). Just as with the measurement of general health, it is important for health care providers to be able to determine ankle health, rather than simply describing it as an absence of disease. The list of functional characteristics identified in this study is an important step in determining how a person with a healthy ankle is different from someone with an unhealthy ankle.

A potential issue with the functional characteristics, though, is that the items on the list are not mutually inclusive. For instance, "patients can be functional without full range of motion," or "they can have pain and still be functional," as also pointed out by one of the experts in their comments. As in the definition of a healthy ankle, some items are not dependent only on ankle function (i.e., gait) and may be limited due to injury to another part of the body. An appropriate addition to each of the characteristics might be "as related to the ankle" to clearly identify the functional ability of the ankle.

The last characteristic: "It depends on what they want to do (it is based on patient's perspectives and goals)" is a key component to all patient outcomes assessment. The patient's goals and objectives must be specifically discussed in order to effectively monitor and measure their recovery and progression (Jette, 2006; Snyder et al. 2008). This characteristic points to the fact that all patients are different and that no two patients will have the exact same characteristics, but how the characteristics are measured and assessed should be tailored to each individual. Even though all patients are different, however, it is important to have a general understanding of what constitutes ankle health and this list of characteristics marks another important step in establishing ankle health.

In addition to the characteristics of a healthy ankle, the experts were asked to identify the characteristics of an unhealthy or acutely injured ankle. An interesting difference between the characteristics of a healthy versus an unhealthy ankle is that the characteristics of an unhealthy ankle included many more items overall, but also many more impairments. Although "impairments" are not "functions," there were several impairments that reached consensus to be included as functional characteristics of an unhealthy ankle. Likewise, there were many more overall characteristics that reached consensus for characteristics of an unhealthy ankle than for a healthy ankle. It seems to be easier to identify the deficits or limitations that result due to injury than the characteristics that constitute health.

Other comments and issues that arose pertain to the term "decreased." As in the list of characteristics of a healthy ankle, the reference point for "normal" was not clearly identified. Also, it is not clear if a patient must have all or just a single characteristic on the list to be considered "unhealthy." As pointed out with the previous questions, items such as "decreased confidence" may be due to factors outside of the ankle, so again, an addition may be to specify "as related to the ankle" to each characteristic.

There were many items that reached consensus on this list that were redundant (i.e., loss of range of motion, loss of dorsiflexion, loss of dorsiflexion/plantarflexion; and neuromuscular inhibition, altered neuromuscular drive, reflex inhibition, decreased proprioception, decreased coordination, decreased balance, decreased postural stability). These items can be further reduced to loss of range of motion and altered neuromuscular drive, respectively. Apprehension and decreased confidence can also be reduced to "decreased confidence."

Similarly to the final item on the list of characteristics of a healthy ankle, "Depends on patient's desired or normal level of activity" was also included as a characteristic of an unhealthy ankle. Once again, this points to the concept that every patient is different and their individual goals and desires should be taken into consideration when evaluating outcomes and measuring state of health. For instance, not all patients will be able to perform "back pedals," even in the healthy state, so it would not be practical to use this as a benchmark for health, or lack of health. On the other hand, being able to back pedal is an important skill for some individuals to be functional in their sport. Therefore, it is imperative to assess the patient's previous level of activity and goals in order to determine how the injury has impacted their functional ability.

Assessment Techniques

The third specific aim of this study was to identify outcome instruments used to assess ankle dysfunction following injury. The hypothesis was that a consensus would be reached. The assessment techniques identified in the Round 3 interview were not presented back to the expert panel until after consensus was reached regarding the functional characteristics of healthy and unhealthy ankles because it did not make logical sense to determine the assessment techniques until the experts knew what they would be measuring. Although the assessment techniques still did not perfectly align with the characteristics, the techniques generally matched the characteristics.

One addition to this list suggested by an expert in his comments was to add "can perform without pain" to the assessments. However, as mentioned previously, it is possible for a patient to function at their full capacity, even with pain. Therefore, it does not seem as though pain should be an indicator of "passing" or "completing" a functional assessment. Just as in the previous discussions, assessments of characteristics that are not functions reached consensus on the list of assessment techniques. However, this more closely aligns the assessment techniques with the identified characteristics of healthy and unhealthy ankles.

One expert commented: "What I really like about this is that it seems as though laboratory-oriented, clinician-oriented, patient-oriented measures have emerged as being relevant for assessment. This is a great sign that we're looking for context among the patient, clinician, and researcher!" Similarly, Wiklund (2004) stated that patient-reported outcomes help link the information gathered with clinician-reported outcomes to determine the patient's perceptions of their injury and progress, which adds important information that cannot be collected only using clinician-reported techniques. Other important reasons for clinicians to incorporate patient-reported outcomes along with clinician-reported measures is that patient-reported outcomes facilitate communication between the patient and clinician, and also may reveal patient concerns or conditions that may have gone unnoticed (Valderas et al., 2008). Laboratory-oriented, clinician-oriented, and patient-oriented assessments did, indeed, emerge throughout the experts' responses to the various questions in the present study. As identified in previous research, it is important for clinicians to talk with patients about their perceptions of recovery, in addition to taking clinician-oriented measurements, because there are often discrepancies between the clinician's and the patient's impressions of readiness to return to activity (Larmer, McNair, Smythe, & Williams, 2011).

In comparison to a study which examined recovery of ankle sprains following discharge from emergency departments, the measured ankle outcomes showed some similarities to the assessments identified in the present study (Aiken, Pelland, Brison, Pickett, & Brouwer, 2008). Figure eight measurement of swelling, ankle circumference measurement for swelling, dorsiflexion range of motion, plantarflexion strength, dorsiflexion strength, plantarflexion peak torque and dorsiflexion peak torque were the clinician-measured outcomes assessed, while the Foot and Ankle Outcome Score (FAOS) was the patient-reported outcome measure. Although some of these outcomes are similar to those identified in the current study, there are also differences. No measurements of swelling were included in the final list in the current study (although it was included as a characteristic of an unhealthy ankle), nor did muscle torque or strength measurements. Similar to both studies are measurements of range of motion and patient-reported outcomes, although the FAOS specifically was not included in the final list in the current study.

Definition of Functional Ankle Instability

The final specific aim of this study was to establish a definition of functional ankle instability and to identify the resulting dysfunction, and it was hypothesized that although a consensus would be reached on the definition of FAI, no consensus would be reached regarding the resulting dysfunction. Thus, it was expected that the expert panel would reach consensus regarding the definition of a healthy ankle, and they, in fact, did. It is interesting that the experts listed many components of the definition of FAI after the Round 1 interview, but after two more rounds of feedback, settled on the definition as "a recurrent sense of giving way of the ankle," which only incorporates one component of the many items that were initially listed. Interestingly, this definition is very similar to the initial definition of FAI as described by Freeman (1965), the tendency of the foot to "give way" (p. 669).

Other authors have described functional instability as "the disabling loss of reliable static and dynamic support of a joint" (Vaes et al., 1998, p. 692). Recently, Hertel (2002) described lateral ankle instability as "the existence of an unstable ankle due to lateral ligamentous damage caused by excessive supination or inversion of the rearfoot" (p. 364) and chronic ankle instability as "the occurrence of repetitive bouts of lateral ankle instability, resulting in numerous ankle sprains" (p. 364). In his research, Hertel does not use the term "functional ankle instability," but instead supports a paradigm in which mechanical and functional insufficiencies as sub-components of chronic ankle instability. Tropp (2002) also supports the concept that functional and mechanical instabilities are components of chronic ankle instability, but describes functional ankle instability as "the subjective feeling of ankle instability or recurrent, symptomatic ankle sprains (or both) due to proprioceptive and neuromuscular deficits" (p. 512).

In addition to the definition of functional ankle instability, the experts were also asked to define chronic ankle instability. The proposed definition did not reach 80%

consensus, which was not surprising because some experts indicated that they feel that FAI and CAI are the same thing, while others feel that CAI is an umbrella term under which FAI is a component. This concept of CAI being an umbrella term is supported by the work of Hertel (2002) and Tropp (2002). Because there were discrepancies reported by the experts, it was not surprising consensus was not reached on both FAI and CAI. Additionally, there were very strong comments provided in support of FAI, in contrast to CAI, and vice versa. For instance, in one comment an expert stated in regards to FAI: "....if you must use that term." In another comment a different expert stated CAI: "....is a bad term...." and FAI "....is the better term...." Not only, then, are there disagreements as to the definitions of FAI and CAI, but also as to if either term should be used at all. Therefore, the likelihood that the greater healthcare profession will be satisfied with the proposed definition is minimal.

Just as with healthy and unhealthy ankles, the experts were asked to describe the functional characteristics of a functionally unstable ankle and a chronically unstable ankle. Some of the characteristics included in the final list of function characteristics of a functionally unstable ankle were the same, or similar to, the characteristics of an unhealthy ankle, but there were some differences. As compared to previous reported lists of characteristics that identify FAI (C. Brown, 2011; Demeritt et al., 2002; Hadadi et al., 2010; Hubbard, Kramer, Denegar, & Hertel, 2007a; Knapp et al., 2011; McKnight & Armstrong, 1997; Ross et al., 2009; Van Bergeyk et al., 2001) the list that was agreed upon in the current study is not an exact match with any of them, although some of the components were the same. Interestingly, similar to the characteristics identified for healthy and unhealthy ankles, the last item on this list was "it depends on what they want to do." Once again, this emphasizes the importance of considering the desires and goals of the patient as they return to their previous level of activity.

Another point of interest is that the item "there is a continuum" reached consensus as a characteristic of FAI. This item coincides with comments made by one of the experts that "it is a syndrome" and "not all patients are heterogeneous." Another expert commented that they would have liked to have been able to select "not necessarily." Even though a patient with FAI must not necessarily have *all* of the characteristics agreed upon by the experts, it is important to identify the signs and symptoms that comprise FAI as a syndrome, just as the signs and symptoms that identify other syndromes, such as anterior compartment syndrome or carpal tunnel syndrome. Both of those syndromes have common signs and symptoms that lead to the diagnosis or ruling out of the condition. It seems that the same system of identification should be possible for FAI, but the common characteristics that must be present with FAI must first be determined.

In contrast to FAI, none of the functional characteristics of CAI reached consensus. Some of the experts stated that the functional characteristics of CAI are the same as FAI, while others stated that it is the same, just a longer period of time. Interestingly, three experts chose "I am not sure if there is a performance decrease or cost" as related to the functional characteristics of CAI. This implies that there are still misconceptions and misunderstandings about the concept of CAI. Because the definition of CAI did not reach consensus, it was not surprising that the functional characteristics did not reach consensus.

Other Suggested Terms

The content experts were asked if there were any other terms that they use in addition to, or in place of FAI or CAI. They were also asked if they felt if "persistent ankle instability" is an appropriate substitute. No items reached consensus as a term used for FAI or CAI, however, terms that emerged from the comments to these questions included "recurrent ankle instability," "recurrent ankle dysfunction," "dysfunction." Some experts felt very strongly that the terms FAI and/or CAI be changed to something different, while others stated that FAI and CAI are interchangeable terms. One expert commented that that no single term will encompass all patients and that clinical prediction rules should be explored, which agrees with the idea the ankle instability in general is a syndrome. Ultimately, however, these results indicate that there are still many inconsistencies and contradictions in regard so the concept of ankle instability.

Summary

In summary, a working definition of a healthy ankle was established, along with the functional characteristics and assessment techniques of healthy and unhealthy ankles. Consensus was also reached regarding the definition of FAI; however, there are still many disagreements, as described in the experts' comments. The definition of CAI did not reach consensus, but the majority of the experts did agree on the definition. Although the experts voiced mixed opinions regarding the definitions of both FAI and CAI, these results offer a good starting point for the working definitions.

CHAPTER IV

STUDY 2

Although ankle injury represents the most common injury evaluated and treated by athletic trainers, there is no consensus as to the functional outcomes that should be measured following injury. Even though many outcomes instruments exist, none of them are appropriate for use with the elite level individuals. Therefore, the purpose of this study was to identify the functional outcomes that should be assessed following ankle injury. Study 1 examined the definitions, characteristics, and assessment techniques of healthy and unhealthy ankles, as well as of functional and chronic ankle instability, from the view point of content experts. This study investigated elite athletes' perceptions of the impact that ankle and foot injuries have on function in daily life, practice, and competition.

Research Design

An observational, non-experimental design was used involving one-on-one interviews. All procedures were carried out in accordance to the University of Northern Iowa's Institutional Review Board policy. Detailed recruitment procedures and scripts, as well as consent documents, are included in the appendices (Appendix B3 – B5). No participants were excluded based on ethnic, racial, religious or cultural backgrounds.

Methods

The purpose of this study was to identify the functional outcomes that should be assessed following ankle injury in individuals with high levels of physical ability.

Specific Aim 1

To determine the functional limitations suffered by athletes following an acute ankle sprain.

Hypothesis 1

The athletes would identify functional limitations that include tasks that are more difficult than the items on the FADI.

Specific Aim 2

To gather injured athletes' feedback regarding the difficulty of the FADI.

Hypothesis 2

The athletes' feedback would indicate that the items are the FADI are too easy.

Participants

Nineteen (10 female, 9 male) athletes were recruited from an NCAA Division IAA institution (n = 15) and a USHL amateur junior ice hockey team (n = 4) who had experienced a foot or ankle injury participated in this study. The athletes were from nine sports including track and field (n = 4), soccer (n = 4), ice hockey (n = 4), wrestling (n = 2), basketball (n = 1), volleyball (n = 1), football (n = 1), tennis (n = 1) and softball (n = 1). Detailed demographics are presented in Table 13. The definition of injury included injuries that had occurred in the past six months to allow for the inclusion of athletes who were in the process of returning to athletic participation following their injury. The athletes' responses to each of the interview questions were analyzed qualitatively and the emergent themes are presented categorically. Direct quotes from the interviews will be used to depict the common themes.

Table 13

Athletes' Demographic Data

Demographic Characteristic		*n*
Sex		
	Male	9
	Female	10
Sport		
	Track and Field	4
	Soccer	4
	Ice Hockey	4
	Wrestling	2
	Basketball	1
	Volleyball	1
	Football	1
	Tennis	1
	Softball	1
Injury type		
	High ankle sprain	2
	Lateral ankle sprain	9
	Foot sprain	1
	Talus fracture	1
	Distalfibula fracture	1
	Malleolar contusion	2
	Laceration	1
	Chronic inflammation	1
	Tendonopathy	1

Other		Range (mean ± SD)
	Age (years)	18 – 22 (19.63 ± 0.9)
	Years exp*	7 – 16 (12.31 ± 3.09)
	Time since injury (wks)	0.57 – 20 (5.43 ± 5.79)
	Time to RTP** (days)	3 – 14 (7.4 ± 3.6)
	Was not out	*n* = 7
	Have not returned	*n* = 5
FADI Scores		
	ADL	27 – 104 (79.58 ± 24.24)
	Sport	0 – 32 (18.32 ± 11.09)
	Total	27 – 136 (97.89 ± 34.39)

*Exp = Experience, **RTP = return to play

Instruments

Each athlete completed an injury history form (Appendix B1), along with the FADI (Martin et al., 1999; Appendix B2). The injury history form asked for participant demographics including height, weight, age, sex, injury status, and the presence of ankle pathology. The FADI consists of 34 items which assess function and pain with activities of daily living and sport activities. The FADI consists of 34 items that are separated in to the FADI-ADL and the FADI-Sport. The items are scored on a 0 – 5 scale. For most of the items a score of 0 indicates "unable to do" and a score of 4 indicates "no difficulty." For the pain items, however, 0 indicates "none" and a score of 4 indicates "unbearable." The total possible points for the FADI-ADL are 104; total possible points for the FADI-Sport are 32. The interview questions (Appendix B3) were read aloud to the athletes and their responses, with permission, were recorded using an Olympus WS-210S Digital Voice Recorder (3500 Corporate Parkway, Center Valley, PA). The interview questions reflected on the FADI as it pertained to the athletes' level of function. The subsequent questions addressed how the athlete's injury impacted their everyday life, athletic performance, desired level of function, and peak performance indicators. The final questions addressed the most physically challenging tasks that they undertake during their participation and if/how their injury affected their performance of those tasks.

Procedures

Prior to the interview, each participant provided informed consent in accordance to the University of Northern Iowa Institutional Review Board's requirements. To ensure that the participants understood the purpose of the study, their role in the study, and the procedures of the interview, they were oriented at the beginning of the interview (Appendix B3). Following the orientation, the participants were provided the opportunity to ask questions regarding the study. After completing the history form and the FADI, they were asked to discuss their experiences following ankle injury and their perceptions of the appropriateness of the items on the FADI. After each response, the participants were asked if their answer was complete before proceeding to the next question.

Interview components. Responses were collected from each participant in the following content areas:

1. Feedback regarding the FADI
2. Level of function in everyday life and sport participation
3. Activity limitations caused by the ankle injury
4. Most difficult physical tasks and peak performance markers/indicators

Once all of these components were completed, the participant's involvement in the study was complete. However, with the participants' permission, they were contacted again if follow-up questions arose.

Data analysis. The data collected during the interviews were analyzed descriptively by the primary investigator. The responses were compiled and reduced qualitatively to detect common themes and to categorize the data. Accuracy checks were completed by another experienced researcher to minimize researcher bias and to protect the authenticity of the data. The FADI scores were summed and reported as FADI-Total, FADI-ADL, and FADI-Sport scores.

Results and Discussion

The results of this study will be presented in the order in which the questions were asked during the interviews. Comparisons will be made across the various sports within the results. The discussion will be presented immediately following the results within each section.

Impact of Injury on Everyday Life

The athletes were asked: "How does/did your injury impact your level of function in everyday life?" Responses revealed that the level of impact that ankle or foot injury imposed on the athletes' ability to function in everyday life depended on weight-bearing status. Those who were non-weight-bearing, either at the time of interview or upon initial injury, reported the greatest impact on their everyday life. The football player reported:

> I skip meals because I don't want to get out of bed…it basically confines me to the house…I don't like relying on other people to take care of me…I cook sitting down in a chair…I've been skipping a lot of classes lately because I don't like to ask people for rides.

This athlete's social aversion due to his injury supports previous findings that two people who experience the same injury may suffer vastly different disabilities (Snyder et al., 2008). Similar to this athlete, other athletes who were non-weight-bearing or required the use of crutches reported having difficulties carrying their trays in the student dining centers, climbing into their lofts, walking to class, simply being able to "get up and run somewhere in a hurry."

Athletes who were never non-weight-bearing reported "walking quickly," "running," "pivoting," "going up and down stairs," "jumping," "ice skating," "playing basketball, dodgeball, and football as hobbies," and "not being able to wear high-heeled shoes" as limitations in everyday life. Some athletes, however, did not experience any limitations in their everyday lives. There were no common themes identified among athletes who played the same sport in terms of the impact of injury on everyday life.

It was expected that athletes who were non-weight-bearing experienced the greatest impact in terms of their everyday lives. Some athletes noted how they took for granted the ability to do "everyday" things such as "run where ever I want" and were frustrated when these abilities were taken from them. The football athlete described the most substantial impact on his everyday life. In his comments, he identified social and academic consequences of his injury. He did not like to go out in public because of embarrassment and he was skipping classes because he did not like to ask people for help with transportation to class. Monitoring the social effects of injury on patients is an important component of monitoring patient outcomes (Snyder et al., 2008), but unfortunately clinicians often focus only on the impact injury has on sport and competition. It is imperative that the patient's ability to participate in social engagements is also monitored.

Impact of Injury on Ability to Function in Practice and Competition

Next, the athletes were asked: "How does/did your injury impact your level of function in practice?" and "How does/did your injury impact your level of function in competition?" The limitations that the athletes reported during practice were mostly the same as during competition, so these results will be presented together in this section. Similarly to the impact on their everyday lives, the level of impact that the injury had on the athletes' level of function in practice depended greatly on weight-bearing status. Those who were non-weight-bearing reported much greater functional limitations than those who were not. Some athletes were completely removed from participation, while others experienced no limitations in practice due to their ankle or foot injury. Commonalities emerged within the various sports. Soccer athletes reported several limitations during practice, including: "kicking the ball," "cutting," "planting my foot," "changing directions," and one soccer athlete noted that she was more careful when going into tackles after the injury first happened. Hockey players noted that "pushing off" the injured side, "starting," "stopping," and skating in general were limitations they faced during practice. Additionally, one hockey athlete reported "accelerating, crossing-over, pivoting forward and backward" on skates was impossible for two weeks. Although there were two wrestlers, comparisons could not be made because one of them was not participating at the time of interview due to his injury. Furthermore, even though there were four track and field athletes who were interviewed, none of them participated in the exact same events; therefore, their responses did not have as much overlap as the other sports.

There were also variations in the functional limitations between each of the sports. While the wrestler reported having problems wrestling "in bottom position," the tennis athlete reported "I just do ground strokes while standing in place" (rather than taking short balls which require stepping to the ball), and the basketball player noted "not being able to go up for a lay-up in one fluid motion from a fast-break" as a functional limitation. Additionally, a volleyball player reported not being able to work on "mental toughness" because her injury-modified workouts were not at a high enough intensity to challenge her mental toughness.

A commonality across several sports was that the athletes noted that they did not notice their injury or pain as much during competition because of the "adrenaline" or because they were in "game mode." Contrastingly, one athlete noted that she is more aware of her injury during competition, which limits the fluidity of her motions and ultimately, her functional ability. Athletes across two sports noted that they are more confident during practice and competition while being taped, although one of them noted that it is "probably a mental thing." Another commonality was that athletes reported having pain while practicing or competing, but that "playing through the pain" is "just part of it" or "I don't want to sit out" and "I push until I can't push anymore, even though it hurts." Some athletes in various sports were not participating in practice or competition at the time of the interview, but athletes from multiple sports noted that they would participate if they were given medical clearance to do so (Table 14).

The mentality of "playing through pain," and pain is "just part of it" is consistent with the sport ethic, which identifies "what it means to be a real athlete" (Hughes & Coakley, 1991, p. 308). The sport ethic has been widely accepted by coaches, parents, the media, fans, as well as athletes, which has often led to "overconformity" to the sport ethic by athletes. The sport ethic identifies that playing through pain is as an expected component of being an athlete (Hughes & Coakley, 1991). Therefore, when athletes "overconform" to the sport ethic, they tend to avoid medical care for pain and continue sport participation despite their pain (Hughes & Coakley, 1991).

Additionally, Malcolm (2006) found that adolescent female softball players learn to "abide by" the sport ethic at an early age. Even the athletes who entered the sport with traditionally feminine perceptions of pain adapted to the sport ethic by ignoring pain and playing with injuries (Malcolm, 2006). This finding is congruent with the athletes' comments in regards to pain. At the elite level, it seems to be commonplace for athletes to continue playing, even with pain and injury, regardless of their gender, as both male and female athletes reported playing with pain.

Recently, the measurement of health-related quality of life has been encouraged in the field of athletic training (Valovich-McLeod et al., 2008). Health-related quality of life is a general assessment of how general health status affects a patient's overall quality of life. Although elite athletes' general quality of life is affected during the acute phase of injury, and even more so if placed in non-weight-bearing, this measurement does not seem valuable as the athletes return to participation at the end phases of injury. For instance, as the elite athletes progress to the final stages of recovery and return to previous levels of participation, there are elite-level skills that they still feel limited from, as discovered in the current study. An example of an elite-level skill noted by the basketball player is: "being able to go up for a lay-up in one fluid motion from a fast-break." Therefore, measuring health-related quality of life during this phase of their recovery will not detect any limitations or concerns with participation, which emphasizes the need for instruments that are both difficult and specific to the physical tasks and demands of elite level athletes.

Table 14

FADI Score, % of Return to Activity, Ability to Compete

Athlete	FADI ADL Score	FADI Sport Score	FADI Total Score	Weight-Bearing Status	% of Return	Could you compete?	Would you compete?	Modified practices due to injury?
Soccer 1	94	28	122	WB	85	Yes	Yes	No
Soccer 2	98	31	129	WB	100	Yes	Yes	No
Soccer 3	104	31	135	WB	95	Yes	Yes	No
Soccer 4	84	3	87	WB	40 - 50	No	No	No practice
Wrestling 1	89	20	109	WB	Unsure, <100	Yes	Yes	No
Wrestling 2	27	0	27	NWB	40	If allowed	If allowed	Yes
Hockey 1	103	31	134	WB	100	Yes	Yes	No
Hockey 2	77	20	97	WB	100	Yes	Yes	No
Hockey 3	95	27	122	WB	90	Yes	Yes	No
Hockey 4	104	32	136	WB	100	Yes	Yes	No

(Table continues)

Athlete	FADI ADL Score	FADI Sport Score	FADI Total Score	Weight-Bearing Status	% of Return	Could you compete?	Would you compete?	Modified practices due to injury?
T & F 1	84	12	96	WB	80	Yes	No	Yes
T & F 2	77	13	90	WB	60 – 70	Maybe	Maybe	Yes
T & F 3	87	25	112	WB	75 - 80	Yes	No	Yes
T & F 4	89	22	111	WB	90	Yes	Yes	No
Softball	89	21	110	WB	85 - 90	Yes	Yes	No
Basketball	80	18	98	WB	75	Yes	Yes	No
Tennis	70	13	83	WB	70 - 80	Yes	Yes	Yes
Volleyball	29	1	30	WB	Unsure, <100	If allowed	If allowed	Yes
Football	32	0	32	NWB	30 – 40	No	No	No practice

FADI = Foot and Ankle Disability Index, NWB = Non-weight bearing, WB = Weight bearing

Factors Limiting Ankle and Foot Function

The athletes were asked: "Do you feel as though you have returned to 100% of your desired activity?" If they indicated that they were not at 100%, they were asked: "How would you rate your percent of return to full recovery on a scale of 0 – 100%?" and "What is the major limiting factor?" If they indicated that they had returned to 100%, they were asked: "What do you feel was the last major factor that limited your function?" When asked to discuss the factor that limited their return to 100% of their desired activity, or the last function to return as they were returning to full activity, just over 50% ($n = 10$) of the athletes reported pain as the most limiting factor. Other reported factors included "wear and tear" of the ankle, lack of rest, fractured bone, physician's orders, limited range of motion, lack of strength of the ankle, lack of joint stability, fear of pain and re-injury, non-weight-bearing status, inability to stop and start quickly and make quick turns in the corners (in reference to hockey), inability to rotate ankle and foot, cutting, and re-conditioning.

As noted in Table 14, even though some athletes did not rate themselves at 100% of their desired level of activity, they were still actively participating in their sport. One athlete even stated that "fear of pain and re-injury" is what was limiting her full return to function, even though she was still competing. Multiple athletes commented that even though they were experiencing pain and were less than 100% recovered, they were still participating because "I have to," "you learn to live with the pain," and "I don't want to sit out." It is clearly depicted in these interviews that tolerating and even ignoring pain, which is the body's natural warning mechanism that something is wrong, is expected at the elite level of participation. The body's health and well-being is sacrificed in order to be able to continue competing. This, then, begs the question if athletes are truly being asked about their pain level and if proper measures are being taken to reduce the pain and promote healing and recovery prior to return to participation.

Difficult Physical Tasks

The athletes were asked: "What is the most difficult physical task involving the lower extremity that is required of your sport/conditioning/training that requires use of your foot and ankle?" Their opinions of the most difficult physical tasks involving the lower extremity were not identical among athletes who participate in the same sport, but similarities emerged. The most sport-specific activities included "getting out of the blocks" and "going full speed for 30 meters while wearing spikes," which were noted by track athletes. Soccer players identified tasks that involved the ability to sustain their level of performance of an activity over time. Specific tasks reported by multiple soccer players included: "being physically and mentally fit…being able to still play while tired…being focused while tired," "endurance activities, sprints, shuttle drills", "fitness tests, timed 400s, full field killers, shuttle runs, long sprints…doing them repeatedly." While hockey players also included the ability to sustain a high level of performance over time, their physical tasks were more sport-specific and included "skating" because it "requires so many different muscles and cardio and everything is firing all at once," "keeping your play at a consistently high level throughout the game." One hockey player noted off-ice training as the most difficult component (i.e., plyometrics, jumping, starting, stopping, sprints, cone drills, accelerating, and decelerating). Similarly, a wrestler also identified being able to go “harder, when you get tired you go more, you just don’t stop” as the most difficult task, while the other wrestler commented that being on his feet and being in his stance throughout an entire match is challenging because “half a match takes place on your feet.” Athletes across multiple sports identified conditioning-related activities, such as sprint drills, 300 yard shuttles, running stairs while wearing a weighted vest, squatting weight, running hills, and gut busters (a combination of sprinting, jumping, push-ups, planting, pivoting, and core exercises), as the most physically demanding task of their sport or training.

The athletes noted various ways in which their injury impacted their performance of the task that they thought to be the most difficult, ranging from having no impact to eliminating them from that activity. Other comments included “it’s one more thing to think about, but I keep going,” “it hurts, but it doesn’t stop me,” “it limits my performance of that activity,” “I do a modified activity instead,” “something like this screws you up mentally,” “can’t do block starts,” “I wear tennis shoes instead of spikes,” and “I am hesitant to go all out.” The ways in which their injury impacted their performance of the most difficult task varied both within and between sports.

Peak or best performance indicators included objective measurements such as time, speed, finishing place, weight, zero goals scored on you, points scored, hits thrown, number of take-downs, the score difference; as well as subjective measurements, such as “they don’t stop, keep attacking, don’t stall, keep intensity over time, don’t give up,” “verbal validation” (from coaches), “knowing I am at my top speed,” “knowing you are at your personal best,” and “visual measurements” (of skating ability).

Most of the athletes agreed that there was someone who could perform the task that they identified as the most difficult better than them, but the quantification of how much better ranged from "slightly better" to "a great deal better than me." Likewise, most athletes agreed that a professional or Olympic athlete would be able to perform the task better than them. Hockey players and the football player, however, stated that they feel they are at the same level or are near the same level as professional athletes. Most of the hockey players aspired to enter the National Hockey League. With no exceptions, the athletes agreed that their level of performance of the task was better as compared to a high school athlete. It was noted by multiple athletes that the transition from high school to intercollegiate or junior league competition was where the biggest differences are seen.

When asked if their ability or willingness to perform the task was different during practice than during competition, some responded that there was no difference "you train as hard as you compete," and "practice like you're going to start." Others said that during competition they felt a "mental or adrenaline boost during games," or had the "game mentality," or were "distracted by the fact you're in a game" so they felt they worked harder during competition; while others commented that they "strive harder during games," or "would rather go all out during competition." In some cases, this question was inapplicable because the task they identified was not performed during competition.

This portion of the interviews identified many activities and tasks that are required of elite level athletes that are not identified on the FADI. The high-level skills of these athletes are not assessed on that instrument such as sprinting, running for long periods of time, stopping and starting quickly while ice skating, wrestling in bottom position, etc. The most difficult items on the FADI-Sport subscale include running, jumping, cutting/lateral movements, squatting and stopping quickly. Therefore, the FADI is missing the top end of the skills that athletes are inhibited from due to injury. Also evident from the athletes' feedback are the differences that exist between the various sports. Athletes commonly identified tasks that are specific to their sport, which supports the idea that the FADI needs more difficult items that are sport-specific. One way to make this feasible is to transition to a computer adaptive model to accurately assess patient-reported outcomes to enable every item on the outcomes instrument to be specific to each individual's sport or activity.

Computer adaptive tests (CAT) are broad patient-reported outcomes, meaning that they can be used with various populations. Whereas most patient-reported outcome instruments are "static," meaning that there are a specified number of questions and the patient is asked to answer all of them, computer-adaptive tests are "dynamic" (Snyder et al., 2008), meaning that the questions presented to the patient depend upon the response to each subsequent question. Therefore, the CAT is able to self-adjust the questions to precisely measure the patient's ability. Another advantage of CAT is that a large amount of items can be stored and used only when applicable according to the patient characteristics (Helbostad et al., 2009). Ideally, there will be CAT options available to assess patient-reported outcomes of all parts of the body and for various pathologies, but for now we must primarily rely on traditional, static instruments. Transitioning to a computer adaptive model should be considered in the further development of patient-reported outcomes assessment.

Foot and Ankle Disability Index Comments and Modifications

The FADI-Total scores ranged from 27 to 136 of a possible total of 136. The sport sub-scale scores ranged from 0 to 32 of a possible total of 32, while the ADL sub-scale scores ranged from 27 to 104 of a possible total of 104. The only athlete that scored 100% on the FADI was the athlete who had the longest recovery time (20 weeks). Two other athletes were close to maximizing their score on the FADI-Total (134 and 135 out of 136). Their injuries had taken place approximately 12 and 2.5 weeks prior to the interview date, respectively. The athletes with the lowest FADI-Total scores (27 and 32 out of 136) were non-weight-bearing at the time of the interview. As depicted in Table 14, the FADI scores and the athletes' subjective reports of return to participation tended to correspond with one another. Interestingly, though, there were some athletes who rated their return to activity as 100%, but did not achieve a FADI score of 100%. After reviewing the interview transcriptions, these results tended to reflect the "playing through the pain" is "just part of it" and "I don't want to sit out" and "I push until I can't push anymore, even though it hurts" comments. Even though these athletes were experiencing some physical limitations, they reported their return as 100% because they were participating in practice and competitions as they normally would without injury.

The first question that the athletes were asked during the interview, which took place immediately following completion of the FADI, was: "Was there anything missing from this questionnaire?" Upon initial contemplation, most athletes stated that they thought the questions on the FADI were "pretty good," and "they covered everything." In general, the athletes felt that the FADI captured their ability. However, at the end of the interview when asked: "How would you ask questions that are relevant and specific to you in a way that would make sense;" they provided several suggestions. The athletes' responses to both the initial question regarding the FADI, and the final question asked at the end of the interview are presented in Table 15. When asked: "How do you feel about the response choices," most athletes indicated that the choices were "good," "really helpful," "very understandable," or "perfect." Suggestions that were made included adding "not allowed to do" and "haven't tried" as choices, or asking about "pain associated with activity instead of difficulty," and to "make (the responses) more specific to question (time, duration, etc)."

Even though many of the athletes initially responded that the FADI was "thorough" and "captured my ability," they offered a plethora of ways in which the instrument could be made more applicable to the requirements of their sport at the end of the interview. After talking about the ways in which their injury limited them from various activities and functions throughout the interview, they were able to provide many ways in which the FADI could be more applicable to their limitations. As there was in the previous question regarding the most difficult physical task, there was again a resounding recommendation to make the items on the FADI "more sport-specific." Although the athletes agreed that the basic requirements of their sport are represented on the instrument (i.e. running, cutting, jumping), the items are not specific enough. For instance, it was pointed out that "running" is a very general activity and that it would be helpful to include running intensities (jogging and sprinting) as well as distances and/or durations. The same suggestion was made for many of the tasks on the FADI. These findings are in contrast with Finch et al. (2002), who stated that an outcomes instrument must have the ability to discriminate patients of various ability levels and should be free of ceiling effects, which falsely indicate that that the patient is fully recovered. Although a ceiling effect was not detected with this study because only one athlete achieved a maximum score, it is evident from the athletes' responses that more specific and more difficult items need to be added to the FADI to make it more applicable to the demands of their sport.

While it seems that the FADI is a good basic instrument, it is not specific enough to assess the high level activities required of elite level athletes, which implores the need for computerized adaptive tests to accurately assess the functional consequences of injury across all sports. However, the FADI still shows merit with the elite population because of its sports sub-scale. Even though the Foot and Ankle Ability Measure was created from the FADI, and has been psychometrically tested, it still has not been validated for use with the elite athlete population. Future research should attempt to validate the use of the FADI, with more difficult items added, for use with the elite population.

Summary

In summary, it appears that the functional limitations which arise from ankle and foot injuries in the elite athlete population occur on a spectrum. Some athletes reported dramatic limitations in their social, personal, and competitive lives, while others reported only minor limitations. Weight-bearing status seemed to play a major role in the functional limitations experienced by the athletes. Additionally, athletes who participated in the same sport tended to report similar functional limitations and difficult tasks. Although the athletes initially reported that the FADI was "good," at the end of the interview they provided a vast list of items that could be added to the FADI to make it more applicable and relevant to the demands of their sport. Overall, the athletes' feedback identified that the items on the FADI need to be more specific to the requirements of their sport and more difficult to match the physical demands of their sport.

CHAPTER V

CONCLUSIONS, LIMITATIONS, AND FUTURE RESEARCH

Conclusions

The purpose of this research was to develop a model for ankle function which allows researchers and clinicians to better assess patient outcome following ankle injury. Based upon the expert panel's feedback and the comments made by elite athletes, an ankle function model was developed (Figure 1). Whereas the experts identified 16 components of a healthy ankle, from which a definition was derived, the athletes consistently indicated that existing functional measures lacked specificity and were not difficult enough.

The information from the experts regarding a healthy ankle was categorized into five sub-components: function, participation, impairments, pathology, and pain, which were used to create the definition of a healthy ankle. Although not a perfect match, these five categories reflect the components of the WHO-ICF model of health, which provides a common international language to describe disability which spans all realms of health care, and spans time. Although the structure of the WHO-ICF model is quite complex, there are basic terms and concepts. Impairment refers to loss in function or structural abnormalities at the body organ or system level (i.e. dorsiflexion range of motion, tibialis anterior muscle strength; Jette, 1995). Impairments identified by the experts in regards to the 16 components of a healthy ankle included pain, range of motion, strength, artho/osteokinematics, circulation, sensation, and swelling. Over 50% of the athletes identified pain as the major factor that limited their function, which was the primary impairment mentioned by the athletes.

Participation, on the other hand, describes the multifaceted relationships between the pathology from which the person is suffering and the environment in which they function, including social functioning (World Health Organization, 2002). In regards to participation, the experts identified "no restrictions in desired participation" as one of the 16 components of a healthy ankle, while the athletes reported several participation limitations. For instance, some athletes commented that going to class was problematic because they had difficulty walking or were non-weight bearing. Others reported difficulties with self-care, cooking, and social interactions. Although the experts did identify participation restrictions as a component of ankle health, the athletes were much more specific about their participation limitations.

That the five sub-categories reflected the components of the WHO-ICF model of health indicates that health care professionals do utilize a somewhat common and consistent approach which matches that of the WHO-ICF. For instance, both impairments and participation were identified by the expert panel, which match the WHO-ICF model. However, there were some differences. The experts noted that functional limitations are dependent on what the patient desires to do (i.e., specificity). From the athletes' comments, it was also inferred that their functional limitations are specific to their sport. However, "specificity" is not a distinct component of the WHO-ICF model, but seems to be relevant to the "participation" component.

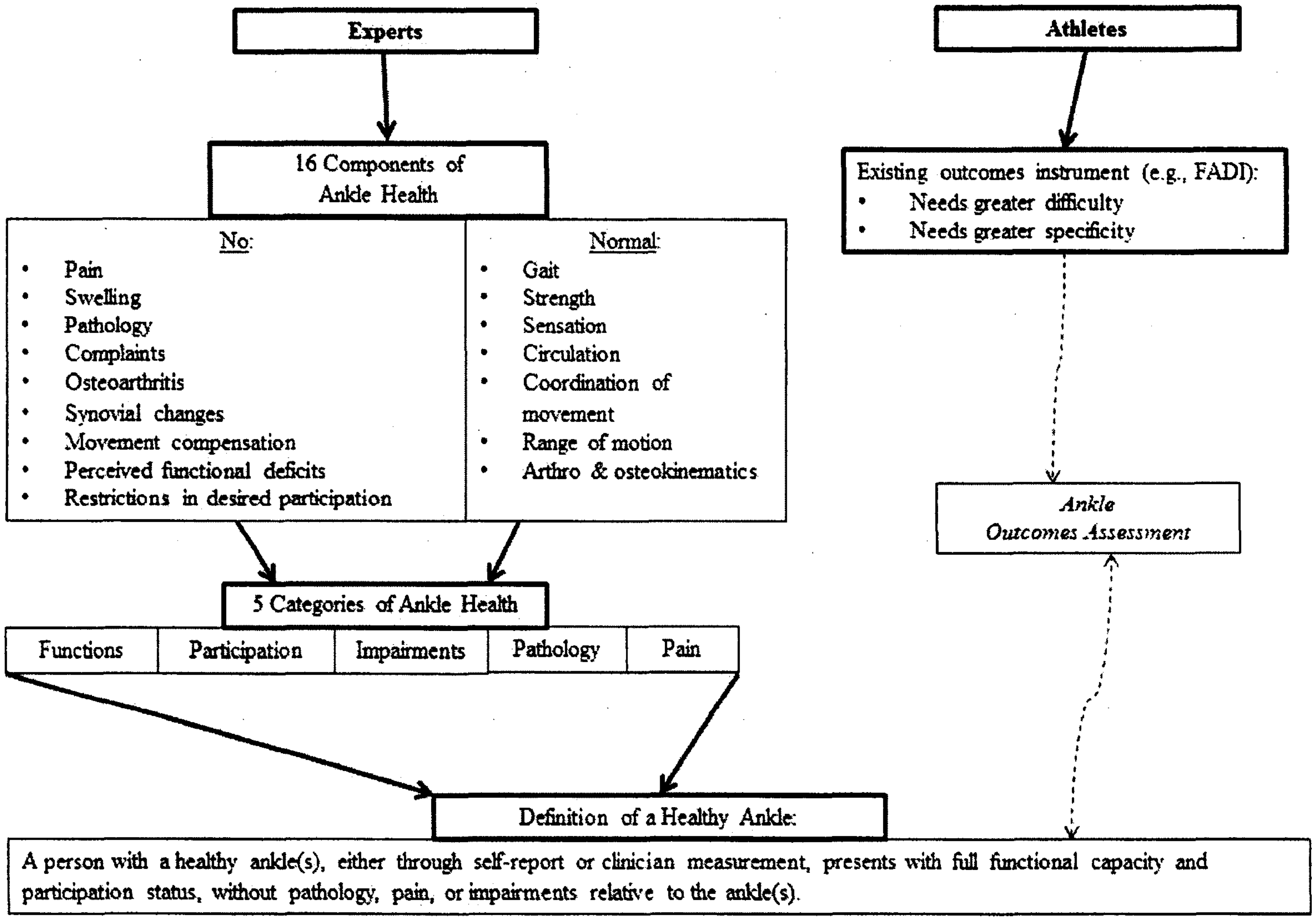

Figure 1. Ankle Function Model

Additionally, the athletes pointed out that items on the self-reported instrument, the FADI, should be more difficult by adding more intensity and duration-specific items. The experts were not asked to comment on difficulty specifically and most of the items they identified did not include intensity or duration. Difficulty is not a component of the WHO-ICF model and does not seem to fit in their model. This should be explored in future research by interviewing a larger group of athletes about the most difficult physical tasks required of their sport and the importance of measuring difficulty.

The input from the experts and athletes delineates that outcome instruments, such as the FADI, need to include items that are relevant to the patient's desired level of function, are specific to their sport or desired activity, and are difficult enough to accurately assess the functional limitations of individuals with high levels of physical ability. As indicated by both the experts and athletes, items on outcomes instruments must be specific to the patient's desired level of activity. Based upon the feedback from the athletes, the instrument could become more specific and difficult by adding items that are more intensity-specific (i.e., running, jogging, sprinting), duration-specific by providing time increments, pain-specific (i.e., can perform without pain), use of external support (i.e., taping and/or bracing), and footwear requirements (tennis shoes, spikes, skates, etc.).

Not surprisingly, pain was identified by both the experts and the athletes. The research committee felt it was important enough to include in the definition even though pain is, in fact, an impairment, which was also a term included in the definition. When addressing pain, 14 of 16 experts indicated that it does belong in the definition of a healthy ankle. Those who disagreed did so because pain is classified as an impairment (World Health Organization, 2002) and, therefore, was redundant to include the impairment "pain" in the definition along with the word "impairment". The justification for leaving the term "pain" in the definition was that the WHO-ICF language is not yet universal to all health care providers. Those who are not familiar with this language may not appreciate that pain is universally included in the definition with the inclusion of the term "impairment." The interesting point is not that both experts and athletes agreed that pain is an important component of injury, but it was how each described pain in relation to function. Experts simply stated that pain is a component of an unhealthy/acutely injured ankle, while the athletes' comments regarding pain were much more specific. For instance, over half of the elite athletes identified pain as the major limiting factor in their return to full participation. Both the experts and athletes noted that pain is part of the injury process, but the experts did not identify pain as a component of a healthy ankle. On the other hand, some athletes reported having pain, but continued to participate in their sport. This begs the question if health care providers are properly assessing pain through self-report. Essentially, there is no other method of measuring pain than by asking the patient, which reinforces the importance of gathering self-reported health status.

Another way in which the experts' and athletes' feedback can be used is to guide how health care professionals assess foot and ankle function. As identified by the experts, subjective tests (i.e., gait analysis, ligamentous tests) along with subjective report from the patient should be used. Likewise, the athletes' comments indicated that it is important to talk to the patient about their level of pain and perception of their recovery. As indicated by Larmer et al. (2011), the clinician's and patient's perceptions of recovery are not the same, so it is important to ask the patient about their readiness to return to activity.

The 16 components of a healthy ankle, as identified by the content experts, were not all reflected on the FADI, which indicates that there are discrepancies between the items on the FADI and how the experts defined ankle health. The FADI does not include items regarding arthro/osteokinematics, circulation, sensation, synovial changes, or osteoarthritis; nor did the athletes mention these components. The discrepancies revealed between the FADI, the athletes, and the experts suggests that there are discrepancies between what clinicians determine to be components of health, what athletes feel are consequences of injury, and the items on the FADI. It seems as though there should be alignment of what is measured by the clinician, the patient's perceptions of the consequences of injury, and items on self-reported instruments, before patient outcomes can be accurately assessed.

Interestingly, although "a recurrent sense of giving way" was identified by the experts as the definition of FAI, none of the athletes with ankle injuries commented that this was a functional limitation. Furthermore, the experts did not establish this as a characteristic of a healthy or unhealthy ankle, which suggests that the entity of FAI may not have functional consequences. Thus, in terms of enhancing our knowledge regarding FAI and CAI, this study did not offer any advancement. Although a consensus was reached regarding the definition of FAI, it is nearly the exact same definition as the original definition offered by Freeman in 1965. No consensus was reached regarding the definition of CAI. Therefore, it seems that the knowledge of FAI and is circular in nature, rather than progressive. Furthermore, because there were strong comments provided by the experts in terms of both FAI and CAI, it was not surprising that advancements were not made regarding these entities. It does not seem likely that the results of this study will encourage researchers to change their definition of FAI or CAI, or their perceptions of the functional cost of FAI/CAI. However, experts should heed the results of this study, which indicate that elite athletes with ankle and foot injury do not complain about a recurrent sense of giving way of their ankle(s). Which, then, begs the question if FAI/CAI are truly problematic in terms of functional ability. If they are not, then researchers should be cautious about the resources put forth in studying these entities.

Limitations and Future Research

The Delphi method is not a flawless research method, but was appropriate for the purposes of this research. Although a panel of 16 experts participated, there are still many expert opinions that are not voiced in this study. Even though a sample was recruited which included experts from various backgrounds, personal bias can still have an effect on the results. The questions asked to the experts could be presented to a larger group of experts to determine if the results differ, then a smaller focus group could be convened to provide any additional feedback and final modifications.

Another limitation could lie within the sample of athletes that were interviewed. Although elite athletes from nine sports participated, the results are not representative of all types of athletes. For example, there are 22 NCAA sanctioned sports (sports that have both men's and women's teams were combined), which left 13 sports unrepresented. Future studies should include athletes from additional sports to evaluate their functional limitations due to ankle and foot injury, as well as their general functional requirements. More athletes from the sports included in this study should be interviewed since in five of the nine sports, only one athlete was recruited.

Because the athletes mentioned several physical tasks that are beyond the physical skill level required of the items on the FADI, those tasks should be further examined and used to create new, more difficult items to be added to an existing outcomes instrument, such as the FADI. This new instrument should then be distributed to elite level athletes from multiple sports to obtain their impressions of the instrument's relevance and applicability to their sport. The instrument should then undergo psychometric analysis to determine its validity with the elite population. Additionally, computer adaptive testing provides an efficient method of measuring outcomes which are specific to the demands of the patient's social and competitive life. This method of outcomes assessment should be further explored in future research.

In order for health care professionals to accurately assess outcomes following injury, instruments which include items that are specific to the patient's sport or activity, and are difficult enough to match the demands of the activity, must be created. This is an important step in facilitating the implementation of evidence-based practice in athletic training. Evidence-based practice is the integration of clinician experience, patient values, and research. By interviewing elite athletes and content experts, this study revealed that both clinician-reported and patient-reported outcomes are important, but that the patient-reported outcomes need to be specific to the patient's desired level of activity and type of activity.

As future research continues to develop items that are specific to the patient's desired level of activity and difficult enough for those with high levels of physical ability, perhaps focus groups of both experts and elite athletes should be convened to create appropriate items. Additionally, perhaps the experts and athletes should have the opportunity to provide feedback on one another's comments in order to make adjustments where necessary. Perhaps seeing each other's comments will elicit new thoughts and ideas and will help to mesh the opinions of both groups to create items that will provide the most insight regarding patient recovery.

By creating a model of ankle health, and identifying the functional limitations that elite level athletes suffer due to foot and ankle injury, this study provides the framework for establishing accurate outcomes assessment of individuals with high levels physical ability, and thereby facilitates the implementation of evidence-based practice in athletic training.

REFERENCES

Agel, J., Beskin, J., Brage, M., Guyton, G., Kadel, N., Saltzman, C. L., ...Thordarson, D. B. (2005). Reliability of the Foot Function Index: A report of the AOFAS Outcomes Committee. *Foot & Ankle International, 26*, 962-967.

Aiken, A. B., Pelland, L., Brison, R., Pickett, W., & Brouwer, B. (2008). Short-term natural recovery of ankle sprains following discharge from emergency departments. *Journal of Orthopaedic and Sports Physical Therapy, 38*, 566-571.

American Education Research Association, America Psychological Association, & National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

Aubry, M., Cantu, R., Dvorak, J., Graf-Baumann, T., Johnston, K., Kelly, J., ...Schamasch, P. (2002). Summary and agreement statement of the first International Conference on Concussion in Sport, Vienna 2001. *British Journal of Sports Medicine, 36*, 6-10.

Baylor, C., Hula, W., Donovan, N., Doyle, P., Kendall, D., & Yorkston, K. (2011). An introduction to item response theory and Rasch models for speech-language pathologists. *America Journal of Speech-Language Pathology, 20*, 243-259.

Bellamy, N. (2005). Science of assessment. *Annals of the Rheumatic Diseases, 64 (suppl)*, 1142-1145.

Beurskens, A. J. H. M., de Vet, H. C. W., & Koke, A. J. A. (1996). Responsiveness of functional status in low back pain: A comparison of different instruments. *Pain, 65*, 71-76.

Borowski, L. A., Yard, E. E., Fields, S. K., & Comstock, R. D. (2008). The epidemiology of US high school basketball injuries, 2005-2007. *The American Journal of Sports Medicine, 36*, 2328-2335.

Bottomley, A., Jones, D., & Claassens, L. (2009). Patient-reported outcomes: Assessment and current perspectives of the guidelines of the Food and Drug Administration and the reflection paper of the European Medicines Agency. *European Journal of Cancer, 45*, 347-353.

Boulkedid, R., Abdoul, H., Loustau, M., Sibony, O., & Alberti, C. (2011). Using and reporting the Delphi method for selecting healthcare quality indicators: A systematic review. *PLoS One, 6* (6).

Brodysky, A. R., O'Malley, M. J., Bohne, W., H., Deland, J. A., & Kennedy, J. G. (2005). An analysis of outcome measures following the brostrom-gould procedure for chronic lateral ankle instability. *Foot & Ankle International, 26*, 816-819.

Brown, C. (2011). Foot clearance in walking and running in individuals with ankle instability. *The American Journal of Medicine*. doi: 10.1177/0363546511408872

Brown, T. (2010). Construct validity: A unitary concept for occupational therapy assessment and measurement. *Hong Kong Journal of Occupational Therapy, 20*, 30-42.

Budiman-Mak, E., Conrad, K., & Roach, K. (1991). The foot function index: A measure of foot pain and disability. *Journal of Clinical Epidemiology, 44*, 561-570.

Button, G., & Pinney, S. (2004). A meta-analysis of outcome rating scales in foot and ankle surgery: Is there a valid, reliable, and responsive system? *Foot & Ankle International, 25*, 521-525.

Carcia, C. R., Martin, R. L., & Drouin, J. M. (2008). Validity of the foot and ankle ability measure in athletes with chronic ankle instability. *Journal of Athletic Training, 43*, 179-183.

Chung, V., Ma, P., Lau, C., & Griffiths, S. (2012). Developing policy for integrating biomedicine and traditional Chinese medical practice using focus groups and the Delphi technique. *Evidence-Based Complementary and Alternative Medicine, 2012*.

Cook, D., & Beckman, T. (2006). Current concepts in validity and reliability for psychometric instruments: Theory and application. *The American Journal of Medicine, 119*, 166.e167-166.e116.

Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science, 9*, 458-467.

De Bie, R., De Vet, H. C., van den Wildenberg, F. A. J. M., Lenssen, T., & Knipschild, P. (1997). The prognosis of ankle sprains. *International Journal of Sports Medicine, 18*, 285-289.

de Vet, H. C. W., Terwee, C. B., & Bouter, L. M. (2003). Clinimetrics and psychometrics: Two side of the same coin. *Journal of Clinical Epidemiology, 56*, 1146-1147.

Delahunt, E., Coughlan, G. F., Caulfield, B., Nightingale, E. J., Lin, C.-W. C., & Hiller, C. E. (2010). Inclusion Criteria When Investigating Insufficiencies in Chronic Ankle Instability. *Medicine & Science in Sports & Exercise, 42*, 2106-2121. doi: 2110.1249/MSS.2100b2013e3181de2107a2108a

Demeritt, K. M., Shultz, S. J., Docherty, C. L., Gansneder, B. M., & Perrin, D. H. (2002). Chronic ankle instability does not affect lower extremity functional performance. *Journal of Athletic Training, 37*, 507-511.

Denegar, C. R., & Miller, S. J. (2002). Can chronic ankle instability be prevented? Rethinking management of lateral ankle sprains. *Journal of Athletic Training, 37*, 430-435.

Denegar, C. R., Vela, L. I., & Evans, T. A. (2008). Evidence-based sports medicine: Outcomes instruments for active populations. *Clinics in Sports Medicine, 27*, 339-351.

DiGiovanni, B. F., Partal, G., & Baumhauer, J. F. (2004). Acute ankle injury and chronic lateral instability in the athlete. *Clinics in Sports Medicine, 23*, 1-19.

Domsic, R., & Saltzman, C. L. (1998). Ankle osteoarthritis scale. *Foot & Ankle International, 19*, 466-471.

Donabedian, A. (1980). *Exploration in Quality Assessment and Monitoring, Vol. 1: The Definition of Quality and Approaches to Its Assessment* (Vol. 1). Ann Arbor: Health Administration Press.

Donabedian, A. (1988). Quality assessment and assurance: Unity of purpose, diversity of means. *Inquiry, 25*, 173-192.

Eberman, L., & Clearly, M. (2011). Development of a heat-illness screening instrument using the Delphi panel technique. *Journal of Athletic Training, 46*, 176-184.

Eechaute, C., Vaes, P., Van Aerschot, L., Asman, S., & Duquet, W. (2007). The clinimetric qualities of patient-assessed instruments for measuring chronic ankle instability: A systematic review. *BMC Musculoskeletal Disorders, 8* (6).

Evans, T. A., & Lam, K. C. (2011). Clinical outcomes assessment in sport rehabilitation. *Journal of Sport Rehabilitation, 20*, 8-16.

Farrugia, P., Goldstein, C., & Petrisor, B. A. (2011). Measuring foot and ankle injury outcomes: Common scales and checklists. *Injury, 42*, 276-280.

Fernandez, W. G., Yard, E. E., & Comstock, R. D. (2007). Epidemiology of lower extremity injuries among U.S. high school athletes. *Academic Emergency Medicine, 14*, 641-645.

Finch, E., Brooks, D., Stratford, P., & Mayo, N. (2002). *Physical Rehabilitation Outcome Measures: A Guide To Enhanced Clinical Decision Making* (2nd ed.). Hamilton, Ontario: Lippincott, Williams & Wilkins.

Freeman, M. A. R. (1965). Instability of the foot affer injuries to the lateral ligament of the ankle. *Journal of Bone Joint Surgery British Volume, 47-B*, 669-677.

Freeman, M. A. R., Dean, M. R. E., & Hanham, I. W. F. (1965). The etiology and prevention of functional instability of the foot. *Journal of Bone Joint Surgery British Volume, 47-B*, 678-685.

Geist, M. R. (2010). Using the Delphi method to engage stakeholders: A comparison of two studies. *Evaluation and Program Planning, 33*, 147-154.

Hadadi, M., Mazaheri, M., Mousavi, M. E., Maroufi, N., Bahramizadeh, M., & Fardipour, S. (2010). Effects of soft and semi-rigid ankle orthoses on postural sway in people with and without functional ankle instability. *Journal of Science and Medicine in Sport*. doi: S1440-2440(10)00916-3 [pii] 10. 1016/j.jsams. 2010.12.004

Hale, S. A., & Hertel, J. (2005). Reliability and sensitivity of the foot and ankle disability index in subjects with chronic ankle instability. *Journal of Athletic Training, 40*, 35-40.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing, 32*, 1008-1015.

Haywood, K. L., Hargreaves, J., & Lamb, S. E. (2004). Multi-item outcome meaures for lateral ligament injury of the ankle: A structured review. *Journal of Evaluation in Clinical Practice, 10*, 339-352.

Helbostad, J., Holen, J., Jordhoy, M., Ringdal, G., Oldervoll, L., & Kaasa, S. (2009). A first step in the development of an international self-report instrument for physical functioning in palliative cancer care: A systematic literature review and expert opinion evaluation study. *Journal of Pain and Symptom Management, 37*, 196-205.

Hertel, J. (2002). Functional anatomy, pathomechanics, and pathophysiology of lateral ankle instability. *Journal of Athletic Training, 37*, 364-375.

Hiller, C. E., Kilbreather, S. L., & Refshauge, K. M. (2011). Chronic ankle instability: Evolution of the model. *Journal of Athletic Training, 46*, 133-141.

Holey, E. A., Feeley, J. L., Dixon, J., & Whittaker, V. J. (2007). An exploration of the use of simple statistics to measure consensus and stability in Delphi studies. *BMC Medical Research Methodology, 7* (52). From: http://www.biomedcentral.com /1471-2288/7/52

Hootman, J. M., Dick, R., & Agel, J. (2007). Epidemiology of collegiate injuries for 15 sports: Summary and recommendations for injury prevention initiatives. *Journal of Athletic Training, 42*, 311-319.

Howe-Weichers, S. (2007). *An examination of group agreement among functional ankle instability participant inclusion criteria* (unpublished Master's Thesis). University of Northern Iowa, Cedar Falls.

Hubbard, T. J. (2008). Ligament laxity following inversion injury with and without chronic ankle instability. *Foot & Ankle International, 29*, 305-311.

Hubbard, T. J., Kramer, L. C., Denegar, C. R., & Hertel, J. (2007a). Contributing factors to chronic ankle instability. *Foot & Ankle International, 28*, 343-354.

Hubbard, T. J., Kramer, L. C., Denegar, C. R., & Hertel, J. (2007b). Correlations among multiple measures of functional and mechanical instability in subjects with chronic ankle instability. *Journal of Athletic Training, 42*(2), 361-366.

Hughes, R. & Coakley, J. (1991). Positive deviance among athletes: The implications of overconformity to the sport ethic. *Sociology of Sport Journal, 8*, 307-325.

Ibrahim, T., Beiri, A., Azzabi, M., Best, A. J., Taylor, G. J., & Menon, D. K. (2007). Reliability and validity of the subjective component of the american orthopaedic foot and ankle society clinical rating scales. *The Journal of Foot and Ankle Surgery, 46*, 65-74.

Jette, A. M. (1995). Outcomes research: Shifting the dominant research paradigm in physical therapy. *Physical Therapy, 75*, 965-970.

Jette, A. M. (2006). Toward a common language for function, disability, and health. *Physical Therarpy, 86*, 726-734.

Keeney, S., Hasson, F., & McKenna, H. P. (2001). A critical review of the Delphi technique as a research methodology for nursing. *International Journal of Nursing Studies, 38*, 195-200.

Kitaoka, H. B., Alexander, I., Adelaar, R., Nunley, J. A., Myerson, M., & Sanders, M. (1994). Clinical rating systems for the ankle-hindfoot, midfoot, hallux, and lesser toes. *Foot & Ankle International, 15*, 349-353.

Knapp, D., Lee, S. Y., Chinn, L., Saliba, S., & Hertel, J. (2011). Differential ability of selected postural-control measures in the prediction of chronic ankle instability status. *Journal of Athletic Training, 46*, 257-262.

Larmer, P., McNair, P., Smythe, L., & Williams, M. (2011). Ankle sprains: Patient perceptions of function and performance of physical tasks. A mixed methods approach. *Disability and Rehabilitation, 33*, 2299-2304.

Lin, C.-W. C., Moseley, A. M., Refshauge, K. M., & Bundy, A. C. (2009). The Lower Extremity Functional Scale has good clinimetric properties in people with ankle fracture. *Physical Therapy, 89*, 580-588.

Maarsingh, O. R., Dros, J., van Weert, H. C., Schellevis, F. G., Bindels, P. J., & van der Horst, H. E. (2009). Development of a diagnostic protocol for dizziness in elderly patients in general practice: A Delphi procedure. *BMC Family Practice, 10* (12). Retrieved from: http://www.biomedcentral.com/1471-2296/10/12

Malcolm, N. L. (2006). "Shaking if off" and "toughing it out" Socialization to pain and injury in girls' softball. *Journal of Contemporary Ethnography, 35*, 495-525.

Martin, R. L., Burdett, R., & Irrgang, J. (1999). Development of the foot and ankle index (FADI) [abstract]. *Journal of Orthopaedic and Sports Physical Therapy, 29*, A32-A33.

Martin, R. L., & Irrgang, J. (2007). A survey of self-reported outcome instruments for the foot and ankle. *Journal of Orthopaedic and Sports Physical Therapy, 37*, 72-84.

Martin, R. L., Irrgang, J., Burdett, R., Conti, S., & Van Swearingen, J. (2005). Evidence of validity for the foot and ankle ability measure (FAAM). *Foot & Ankle International, 26*, 968-983.

McKay, G. D., Goldie, P. A., Payne, W. R., & Oakes, B. W. (2001). Ankle injuries in basketball: Injury rate and risk factors. *British Journal of Sports Medicine, 35*, 103-108.

McKeon, P. O., Ingersoll, C. D., Kerrigan, C. D., Saliba, E., Bennet, B. C., & Hertel, J. (2008). Balance Training Improves Function and Postural Control in Those with Chronic Ankle Instability. *Medicine & Science in Sports & Exercise, 40* (10), doi: 1810-1819 1810.1249/MSS.1810b1013e31817e31810f31892

McKnight, C. M., & Armstrong, C. W. (1997). The role of ankle strength in functional ankle instability. *Journal of Sport Rehabilitation, 6*, 21-29.

McLeod, T. C. V., Snyder, A. R., Parsons, J. T., Bay, C., Michener, L. A., & Sauers, E. L. (2008). Using disablement models and clinical outcomes assessment to enable evidence-based athletic training practice, part II: Clinical outcomes assessment. *Journal of Athletic Training, 43*, 437-445.

Michell, T. B., Ross, S. E., Blackburn, J. T., Hirth, C. J., & Guskiewicz, K. (2006). Functional balance training, with or without exercise sandals, for subjects with stable or unstable ankles. *Journal of Athletic Training, 41*, 393-398.

Nagi, S. (1965). Some conceptual issues in disability and rehabilitation. In M. Sussman (Ed.), *Sociology and Rehabilitation* (pp. 100-113). Washington, DC: American Sociological Association.

National Athletic Trainers' Association Research and Education Foundation. (2011). Retrieved November 20, 2011, from http://www.natafoundation.org/

Olerud, C., & Molander, H. (1984). A scoring scale for symptom evaluation after ankle fracture. *Archives of Orthopaedic and Trauma Surgery, 103*, 190-194.

Panayides, P., Robinson, C., & Tymms, R. (2010). The assessment revolution that has passed England by: Rasch measurement. *British Educational Research Journal, 36*, 611-626.

Parker, J., Nester, C., Long, A., & Barrie, J. (2003). The problem with measuring patient perceptions of outcome with existing outcome measures in foot and ankle surgery. *Foot & Ankle International, 24*, 56-60.

Parsons, J. T., Valovich McLeod, T. C., Snyder, A. R., & Sauers, E. L. (2008). Change is hard: Adopting a disablement model for athletic training. *Journal of Athletic Training, 43*, 446-448.

Pugia, M., Middel, C., Seward, S., Pollock, J., Hall, R., Lowe, L., ...Henderson, N. E. (2001). Comparison of acute swelling and function in subjects with lateral ankle injury. *Journal of Orthopaedic and Sports Physical Therapy, 31*, 384-388.

Revicki, D., Hays, R., Cella, D., & Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology, 61*, 102-109.

Ross, S. E., Guskiewicz, K. M., Gross, M. T., & Yu, B. (2009). Balance measures for discriminating between functionally unstable and stable ankles. *Medicine & Science in Sports & Exercise, 41*, 399-407.

Sackett, D., Rosenberg, W., Gray, J., Haynes, R., & Richardson, W. (1996). Evidence based medicine: What it is and what it isn't. *BMJ, 312*, 71-72.

Sankey, R. A., Brooks, J. H. M., Kemp, S. P. T., & Haddad, F. S. (2008). The Epidemiology of Ankle Injuries in Professional Rugby Union Players. *The American Journal of Sports Medicine, 36*, 2415-2424.

Schlitz, E., Evans, T., & Ragan, B. (2009). *Psychometrics of Ankle Self-Reported Surveys (PASS)* (unpublished Master's Thesis). University of Northern Iowa, Cedar Falls.

Schumacker, R. (2010). Classical test analysis. *Applied Measurement Associates, LLC.* Retrieved November 20, 2011, from: http://appliedmeasurementassociates. com/ama/assets/File/Standards%2520 for%2520Interpreting%2520Reliability% 2520Coefficients.pdf

Schweizer, K. (2011). On the changing role of Chonbach's α in the evaluation of the quality of a measure. *European Journal of Psychological Assessment, 27*, 143-144.

Scranton, P. E. J., McDermott, J. E., & Rogers, J. V. (2000). The relationship between chronic ankle instability and variations in mortise anatomy and impingement spurs. *Foot & Ankle International, 21*, 657-664.

Sefton, J. M., Hicks-Little, C. A., Hubbard, T. J., Clemens, M. G., Yengo, C. M., Koceja, D. M., ...Cordova, M. L. (2008). Segmental spinal reflex adaptations associated with chronic ankle instability. *Archives of Physical Medicine and Rehabilitation, 89*, 1991-1995.

Sefton, J. M., Yarar, C., Hicks-Little, C. A., Berry, J. W., & Cordova, M. L. (2011). Six weeks of balance training improves sensorimotor function in individuals with chronic ankle instability. *Journal of Orthopaedic and Sports Physical Therapy, 41*, 81-89.

Sinha, I. P., Smyth, R. L., & Williamson, P. R. (2011). Using the Delphi technique to determine which outcomes to measure in clinical trials: Recommendations for the future based on a systematic review of existing studies. *PloS Med, 8* (1).

Slattery, E., Voelker, C., Nussenbaum, B., Rich, J., Paniello, R., & Neely, J. (2011). A practical guide to surveys and questionnaires. *America Academy of Otolaryngology-Head and Neck Surgery, 144*, 831-837.

Snyder, A. R., Parsons, J. T., Valovich-McLeod, T. C., Bay, C., Michener, L. A., & Sauers, E. L. (2008). Using disablement models and clinical outcomes assessment to enable evidence-based athletic training practice, part I: Disablement models. *Journal of Athletic Training, 43*, 428-436.

SooHoo, N., Samimi, D., Vyas, R., & Botzler, T. (2006). Evaluation of the validity of the Foot Function Index in measuring outcomes in patients with foot and ankle disorders. *Foot & Ankle International, 27*, 38-42.

Steurer, J. (2011). The Delphi method: An efficient procedure to generate knowledge. *Skeletal Radiology, 40*, 959-961.

Steves, R., & Hootman, J. M. (2004). Evidence-Based Medicine: What is it and how does it apply to athletic training? *Journal of Athletic Training, 39*, 83-87.

Streiner, D., & Norman, G. (1995). *Health measurement scales: A Practical Guide to Their Development and Use* (2nd ed.). Oxford: Oxford University Press.

Thomas, M. (2011). The value of item response theory in clinical assessment: A review. *Assessment, 18*, 291-307.

Thomson, R., Lewalle, P., Sherman, H., Hibbert, P., Runciman, W., & Castro, G. (2009). Towards an international classification for patient safety: A Delphi survey. *International Journal for Quality in Health Care, 21*, 9-17.

Tinsley, H., & Tinsley, D. (1987). Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology, 34*, 414-424.

Tropp, H. (2002). Commentary: Functional ankle instability revisited. *Journal of Athletic Training, 37*, 512-515.

Unick, G., & Stone, S. (2010). State of modern measurement approaches in social work research literature. *Social Work Research, 34*, 94-101.

Vaes, P., Duquet, W., Casteleyn, P., Handelbert, F., & Opdecam, P. (1998). Static and Dynamic Roentgenographic Analysis of Ankle Stability in Braced and Nonbraced Stable and Functionally Unstable Ankles. *The American Journal of Medicine, 26*, 692-701.

Valderas, J., Kotzeva, A., Espallargues, M., Guyatt, G., Ferrans, C., Halyard, M., ... Alonso, J. (2008). The impact of measuring patient-reported outcomes in clinical practice: A systematic review of the literature. *Quality of Life Research, 17*, 179-193.

Valovich-McLeod, T. C., Snyder, A. R., Parsons, J. T., Bay, C., Michener, L. A., & Sauers, E. L. (2008). Using disablement models and clinical outcomes assessment to enable evidence-based athletic training practice, part II: Clinical outcomes assessment. *Journal of Athletic Training, 43*, 437-445.

Van Bergeyk, A. B., Younger, A., & Carson, B. (2001). CT Analysis of Hindfoot Alignment in Chronic Lateral Ankle Instability. *Foot & Ankle International, 23*, 37-42.

Van Der Wees, P., Hendriks, E., Van Beers, H., Van Rijn, R., Dekker, J., & De Bie, R. (2010). Validity and responsiveness of the ankle function score after acute ankle injury. *Scandinavian Journal of Medicine & Science in Sports, 22*, 170-174.

van Rijn, R. M., Willemsen, S. P., Verhagen, A. P., Koes, B. W., & Bierma-Zeinstra, S. M. (2011). Explanatory variables for adult patients' self-reported recovery after acute lateral ankle sprain. *Physical Therapy, 91*, 77-84.

Vergouw, D., Heymans, M., de Vet, H. C., van der Windt, D. A., & van der Horst, H. E. (2011). Prediction of persistent shoulder pain in general practice: Comparing clinical consensus from a Delphi procedure with a statistical scoring system. *BMC Family Practice, 12* (63). Retrieved from http://www.biomedcentral.com/1471-2296/12/63

Ware, J., Jr., Kosinski, M., & Keller, S. (1996). A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care, 34*, 220-233.

Ware, J., Jr., & Sherbourne, C. (1992). The MOS 36-item short form health survery (SF-36). I. Conceptual framework and item selection. *Medical Care, 30*, 473-483.

Wiklund, I. (2004). Assessment of patient-reported outcomes in clinical trials: The example of health-related quality of life. *Fundamental and Clinical Pharmacology, 18*, 351-363.

Wilson, A., Kavanaugh, A., Moher, R., McInroy, M., Gupta, N., Salbach, N., Wright, F. V. (2011). Development and pilot testing of the challenge module: A proposed adjunct to the gross motor function measure for high-functioning children with cerebral palsy. *Physical & Occupational Therapy in Pediatrics, 31*, 135-149.

World Health Organization (2002). *Towards a Common Language for Functioning, Disability and Health: ICF*. Geneva, Switzerland: World Health Organization.

APPENDIX A1

RECRUITMENT SCRIPT

Hello,
My name is Kelli Snyder and I am a faculty member at the University of Northern Iowa. I am writing to gauge your interest in participating as an expert panelist in my research study regarding ankle function. *(You were referred to me by ………)*

The primary purpose of my research project is to develop a model for ankle function that can be used to address outcomes following ankle injury.

I am assembling a panel of 12 to 20 experts and I am inviting you to participate because you have been identified as an expert based upon your clinical or research expertise.

Your participation would involve brief telephone interviews with me, the first lasting approximately 20 minutes and follow-up contacts lasting substantially shorter each time.

If you are interested and would like to hear more about participating, please reply to my email and I will contact you soon to provide details.

I look forward to hearing back from you to discuss my study. I will follow this email with a telephone call within the next 48 hours.

Sincerely,
Kelli Snyder, Athletic Training Division
University of Northern Iowa

APPENDIX A2

ELECTRONIC INFORMED CONSENT

UNIVERSITY OF NORTHERN IOWA
HUMAN PARTICIPANTS REVIEW
ELECTRONIC INFORMED CONSENT

Investigators: Kelli Snyder, MS, ATC, LAT

The Title: *Development of an Ankle Function Model*

Invitation to participate: You are invited to participate in a research project conducted through the University of Northern Iowa. The University requires that you give your signed agreement to participate in this project. The following information is provided to help you make an informed decision about whether or not to participate.

Nature and Purpose: The primary purpose of this research project is to develop a model for ankle function that can be used to address outcomes following ankle injury. I am assembling a panel of 10 to 15 experts regarding the topic of ankle function. I am inviting you to participate because you have been identified as an expert based upon your clinical or research experience. My goal is to reach a consensus from the expert panel regarding several issues related to ankle and ankle function through feedback and discussion.

Explanation of Procedures: If you choose to be in this study, you will be asked to:

1. Review six interview questions regarding ankle injury and function.
2. Participate in an initial phone interview addressing each question (10 – 20 minutes). *Once I have conducted an initial interview with each expert on the panel, I will then compile the data and send it back to you and all of the other experts. At this time you will have the opportunity to make changes.
3. Participate in brief follow up interviews to collect your feedback on the compiled data. I estimate that these interviews will take five minutes or less. I anticipate that after 2-3 rounds, a consensus can be reached. I will end the interviews five, even if a consensus can't be reached.

- Notes:
    - I will record all interviews, but immediately delete your responses as soon as I transcribe them.
    - I estimate that your total interview time will be no longer than 40 minutes combined.
    - The other panel members will NOT see your name or know your identity.

Risks & Benefits: There are no foreseeable risks in participating & your participation may not benefit you.

Confidentiality: The answers you provide will be kept confidential. Your identity will be kept completely anonymous. Special precautions have been established to protect the confidentiality of your responses. Your interview will be deleted once I enter your responses. Other panelists will not know your identity. If my study is published, I will only be identifying averages for my panel's profession(s), years of experience, and setting(s).

Right to Refuse or Withdrawal: Your participation is completely voluntary. You are free to withdraw from participation at any time or to choose not to participate at all.

Questions: If you have any questions about the survey, please contact either Kelli Snyder at (319) 273-7401 or by e-mail at kelli.snyder@uni.edu or Todd Evans at (319) 273-6152 or by e-mail at todd.evans@uni.edu. If we are not available when you call, please leave a message and we will call back. If you have questions about your rights as a participant in this research project, please contact the Institutional Review Board Human Protections Administrator at (319) 273-6148.

Agreement: By responding to this message and replying back to the investigators indicating that I am willing to participate, I am acknowledging that I am fully aware of the nature and extent of my participation in this project as stated above and the possible risks arising from it. I hereby agree to participate in this project. I acknowledge that I am 18 years of age or older.

Thank you!

Kelli Snyder, Athletic Training Division

University of Northern Iowa

APPENDIX A3

TELEPHONE INVITATION SCRIPT

UNIVERSITY OF NORTHERN IOWA

Invitation to participate as an expert panelist.
Telephone Invitation Script

Investigators: Kelli Snyder, MS, ATC, LAT
Project Title: *Development of an Ankle Function Model*

My name is Kelli Snyder,

I am a faculty member at the University of Northern Iowa.

Yesterday, I sent you an email message explaining my study and gauging your interest in participating.

I am assembling a panel of 12 to 20 experts regarding the topic of ankle sprains and/or chronic ankle instability.

I am inviting you to participate because you have been identified as an expert.

The primary purpose of this research project is to develop a model for ankle function that can be used to address outcomes following ankle injury.

I am calling you now to ask if you have any questions about my study or your role as a potential participant.

Are you willing to participate in my study?

If **yes**: "Thank you and provide details for the next step"; if **no**: "Thank you for your time and consideration".

APPENDIX A4

TELEPHONE INTERVIEW QUESTIONS

- Thank you for your participation and time.

- First, I'd like your permission to record this interview. The purpose is to allow the interview process to go faster and your responses can be transcribed. The recordings will be erased once your responses are typed out.

- I will be writing down some of your responses, so at some point I may pause or may have to read back your response to make sure I understand.

- Before we begin I'd like to briefly go over your role and the procedures as part of the training for you as an expert.

- Your role will be to state your opinion about your concept of ankle function.

- At the end of each question I will ask: "Is your answer complete?" before moving on to the next question.

- After I've interviewed all experts, I will condense the responses and send them back to you and the other experts. At that time I'd like for you to review and refine the responses. We will schedule another interview to gather your feedback. Once again the responses will be condensed and sent back to you for your revision.

- Do you have any questions?

Interview questions:

Part 1: This part of the interview will address the concept of a healthy/normal/non-injured ankle as it pertains to an acute injury.

1. How do you define a healthy/normal/non-injured ankle?

Part 2: This part of the interview will address ankle function.

1. Describe the functional characteristics of someone with a healthy/normal/non-injured ankle.
2. Describe the functional characteristics or limitations of someone with an unhealthy/acutely injured ankle.
    - What can't they do?
    - What are the functional costs?

Part 3 will address assessment of function related to ankle health.

1. How do you assess the functional characteristics that you mentioned above?
    - (*For this section, I will mention that I have the characteristics listed and I reflect back on the list when we finish this question)
    - Example:
        1. If "stability": assess anterior translation, using anterior drawer, KT 1000, Lachmans, etc.
        2. If "balance": ability to stand on one leg with eyes close, using force plate, BESS, etc.
        3. ..."SEBT"....etc....:
            - *Can you provide details . . .*
                a. *what characteristic are you measuring ?*
                b. *how do you measure it?*
                c. *What instrument or test?*
                d. *Do you have a reference or source for that test?*
                e. *What do you use the test to measure?*
2. Are there any other characteristics or assessments that you would like to mention?

Part 4 will address functional ankle instability.

1. What is your definition of FAI?
2. Describe the functional characteristics or limitations of someone with FAI.
    - What are the functional costs/consequences?
        1. *If "balance".....is that they only thing that they can't do?*
    - What can't a person with FAI do?
3. How do you assess the functional characteristics that you mentioned above regarding FAI (if any)?
4. Do you have a different title or name that you use for functional ankle instability?
5. Do you feel that there is a name more appropriate than "Functional Ankle Instability"?
    - Did they or did they not mention CAI?

Part 5 will address chronic ankle instability (CAI) *only if they do not mention it.*

1. The term CAI is sometimes used interchangeably with FAI. Do you agree that they are the same thing?
2. What is your definition of CAI?
    - Is it, or how is it, different than FAI (if they don't elaborate)
3. Describe the functional characteristics of someone with CAI.

- What are the functional costs/consequences?
- What are the costs/consequences to one's health?
- What can't a person with CAI do?

4. How do you assess the functional characteristics that you mentioned above regarding CAI (if they are different than above)?
5. Do you have a different title or name that you use to describe chronic instability?
6. Do you feel that there is a name more appropriate than "Chronic Ankle Instability"?

II. **Do you feel that "persistent ankle instability" is an appropriate term?**

APPENDIX B1

INJURY HISTORY FORM

Please do not put your name on this paper.

Ht. ____feet _____inches **Wt.** _______pounds **Age:** _________ **Gender:** M F

1. What NCAA Division 1 Sport do you play? ________________________________

2. Level/year (i.e. freshman, sophomore, etc.) ________________________________

3. Describe your total years of experience in this sport. __________________________

4. Are you currently experiencing any ankle pain or disability? **Y N**

   3a. If so, please describe the injury type________________________________

   3b. Location: **Left Right**

   3c. Date of Injury: ______________days/weeks/years ago **(circle one)**

   3d. Date of Surgery: _______________days/weeks/years ago **(circle one)**

   3e. Please describe any rehabilitation received for this injury

____

5. Have you experienced injury to this ankle prior to the current injury?
   a. if so, please describe the injury type________________________________
   b. Location: **Left Right**
   c. Date of Injury: ______________days/weeks/years ago **(circle one)**
   d. Date of Surgery: _______________days/weeks/years ago **(circle one)**
   e. Please describe any rehabilitation received for this injury

   ____

6. Have you ever experienced your ankle feeling weak, or as if giving out from under you?

   Y N

7. Do you currently have any other injury or condition that limits your activity level? **Y N**

a. If yes, what side is the other injury located? **Left** **Right**

b. Please describe the injury type____

APPENDIX B2

FOOT AND ANKLE DISABILITY INDEX

Instructions: Please answer every question with one response that most closely describes your condition within the past week. Please circle the number that corresponds with your answer to the question stated.

	No difficulty at all	**Slight difficulty**	**Moderate difficulty**	**Extreme difficulty**	**Unable to do**
Standing	4	3	2	1	0
Walking on even ground	4	3	2	1	0
Walking on even ground without shoes	4	3	2	1	0
Walking up hills	4	3	2	1	0
Walking down hills	4	3	2	1	0
Going up stairs	4	3	2	1	0
Going down stairs	4	3	2	1	0
Walking on uneven ground	4	3	2	1	0
Stepping up and down curbs	4	3	2	1	0
Squatting	4	3	2	1	0
Sleeping	4	3	2	1	0
Coming up on your toes	4	3	2	1	0
Walking initially	4	3	2	1	0
Walking 5 minutes or less	4	3	2	1	0

Walking approximately 10 minutes	4	3	2	1	0
Walking 15 minutes or greater	4	3	2	1	0
Home responsibilities	4	3	2	1	0
Activities of daily living	4	3	2	1	0
Personal care	4	3	2	1	0
Light to moderate work (standing, walking)	4	3	2	1	0
Heavy work (push/pulling, climbing, carrying)	4	3	2	1	0
Recreational activities	4	3	2	1	0
Please rate your pain as it relates to your foot or ankle:	**No pain**	**Mild**	**Moderate**	**Severe**	**Unbearable**
General level of pain	4	3	2	1	0
Pain at rest	4	3	2	1	0
Pain during your normal activity	4	3	2	1	0
Pain first thing in the morning	4	3	2	1	0
Because of your foot or ankle pain, how much difficulty do you have with:	**No difficulty at all**	**Slight difficulty**	**Moderate difficulty**	**Extreme difficulty**	**Unable to do**
Running	4	3	2	1	0
Jumping	4	3	2	1	0
Landing	4	3	2	1	0

Squatting and stopping quickly	4	3	2	1	0
Cutting, lateral movements	4	3	2	1	0
Low-impact activities	4	3	2	1	0
Ability to perform activity with your normal technique	4	3	2	1	0
Ability to participate in your desired sport	4	3	2	1	0

APPENDIX B3

INTERVIEW SCRIPT

Development of an Ankle Function Model: Function Following Ankle Injury

Thank you for your participation and time.

□ First, I'd like your permission to record this interview. The purpose is to allow the interview process to go faster and your responses can be transcribed. The recordings will be erased once your responses are typed out.

□ I will be writing down some of your responses, so at some point I may pause or may have to read back your response to make sure I understand.

□ Before we begin I'd like to briefly go over your role and the procedures as part of the training for you as an expert.

□ Your role will be to state your opinion about your concept of ankle function.

□ At the end of each question I will ask: "Is your answer complete before moving on to the next question?"

□ Do you have any questions?

Interview Questions:

1. How has your ankle injury impact your level of function in everyday life?
    a. What does it limit you from doing? What can't you do that you wanted to do?
2. How has your ankle injury impact your level of function in practice?
    a. What does it limit you from doing? What can't you do that you wanted to do?
3. How has your ankle injury impact your level of function in competition?
    a. What does it limit you from doing? What can't you do that you wanted to do?
4. How does it limit your desired level of activity?
    a. What can't you do now that you could do before your ankle injury?
5. Do you feel you have returned to your desired level of activity?
6. What is the major factor that has limited you from your desired level of activity?
7. What do you consider to be the most difficult physical task associated with your sport or conditioning/training in which you are using your ankle/foot? What is your peak or best performance marker of that task (time, distance, weight, or just completion)? How is your injury impacting your performance of that task?
8. Do you mind if I ask the other participants about the tasks that you mentioned?
9. May I contact you in the future if any follow-up questions arise? What is the best way to contact you?

APPENDIX B4

SCRIPT FOR RECRUITMENT

Development of an Ankle Function Model: Function Following Ankle Injury

Hi, my name is Kelli Snyder and I am an athletic training faculty member. I work with (NAME) your athletic trainer. I am in the process of conducting research on ankle injuries and need volunteers to participate in my study.

I think you might qualify for my study because it appears that you have an ankle injury. Your participation would involve a 15 minute interview with me regarding your functional level following ankle injury. My goal is to discover how ankle injury affects the functional ability of highly skilled individuals. All of my participants receive a gift for their participation.

If you are interested, I will take your name and number and contact you to set u a time to meet.

Do you have any questions about my study? Are you interested in participating?

OPTIONS:

If yes: Collect name and contact information

If no: Thank them for their time and wish them well in their recovery.

APPENDIX B5

INFORMED CONSENT

UNIVERSITY OF NORTHERN IOWA
HUMAN PARTICIPANTS REVIEW
INFORMED CONSENT

Investigators: Kelli Snyder, MS, ATC, LAT

Title: *Development of an Ankle Function Model: Function Following Ankle Injury*

Invitation to participate: You are invited to participate in a research project conducted through the University of Northern Iowa. The University requires that you give your signed agreement to participate in this project. The following information is provided to help you make an informed decision about whether or not to participate.

Nature and Purpose: The primary purpose of this research project is to determine the impact of ankle injury on function in elite athletes. I will be interviewing several collegiate athletes with ankle injuries. My goal is to discover how your ankle injury affected your ability to function in athletic activities. I am inviting you to participate because you recently suffered an ankle injury.

Explanation of Procedures: If you choose to be in this study, you will be asked to:

4. Participate in one individual interview with me about how your function has been impacted by your ankle injury.
5. Complete an injury history form.
6. Complete an ankle injury survey.

**Your time commitment will be approximately 20 minutes. Your participation could also include a brief follow-up interview by phone or email if I need clarification to any of your responses. I will ask you about this possibility once we conclude the initial interview.

With your permission, I will record all interviews, but erase your responses as soon as type out your responses.

Risks & Benefits: This study involves minimal risk to you which entails no risks to your physical and mental health beyond those encountered in the normal course of everyday life. For your time and effort, you will receive a $5.00 gift certificate which is redeemable at any Hy-Vee Gas or Food Store.

Confidentiality: The answers you provide will be kept confidential. Your identity will be kept completely anonymous. Special precautions have been established to protect the confidentiality of your responses. Your interview will be deleted once I enter your responses. If my study is published, I will only be identifying averages for my participant's sport(s), years of participation, gender, and age.

Right to Refuse or Withdrawal: Your participation is completely voluntary. You are free to withdraw from participation at any time or to choose not to participate at all.

Questions: If you have any questions about the survey, please contact either Kelli Snyder at (319) 273-7401 or by e-mail at kelli.snyder@uni.edu or Todd Evans at (319) 273-6152 or by e-mail at todd.evans@uni.edu. If we are not available when you call, please leave a message and we will call back. If you have questions about your rights as a participant in this research project, please contact the Institutional Review Board Human Protections Administrator at (319) 273-6148.

Agreement: I am fully aware of the nature and extent of my participation in this project as stated above and the possible risks arising from it. I hereby agree to participate in this project. I acknowledge that I have received a copy of this consent statement. I am 18 years of age or older.

Thank you!
Kelli R. Snyder, MS, ATC, LAT

(Signature of participant)	(Date)	(Signature of investigator)	(Date)
(Printed name of participant)	(Date)	(Instructor/advisor)	(Date)

[NOTE THAT ONE COPY OF THE ENTIRE CONSENT DOCUMENT (NOT JUST THE AGREEMENT STATEMENT) MUST BE RETURNED TO THE PI AND ANOTHER PROVIDED TO THE PARTICIPANT. SIGNED CONSENT FORMS MUST BE MAINTAINED FOR INSPECTION FOR AT LEAST 3 YEARS]

APPENDIX C

STUDY 1 EXPERT WRITTEN COMMENTS

Study 1, Round 2 Expert Written Comments

Question	Comments
How do you define a healthy/normal/non-injured ankle?	Joint mechanics and arthrokinematics seem redundant and perhaps normal joint position would be redundant with these, too. Normal joint stability and no mechanical laxity seem redundant. Patient can complete full functional assessment and patient can perform deep body squats seem redundant. No complaints seems to capture many of the items - like no complaints of instability, no pain, no perceived functional deficits, no restrictions in desired participation - but the "no complaints" option gives very little information (other choices seem better).
	I feel there is a difference between healthy/normal and non-injured. Depends how you are thinking about non-injured. I interpreted this to mean not currently injured when answering the questions.
	There are some that are not necessarily limited to ankle problems (gait, inability to perform squat, etc), and /or may not relate to nml ankle function (laxity, prior injury)
	"No complaints" is redundant with several items that list specific complaints like "pain", "complaints of instability", "perceived functional deficits", "restriction in desired participation" Also, "patient can complete "full functional assessment" is vague. Just because they can complete it doesn't mean that they pass it. Items like "deep body squats", "no compensations in movement activities", and "normal coordination" are somewhat redundant, but more specific than the "full functional assessment" item.

	The final answer requires more than just ankle health. In other words if you have a normal ankle you may still not be able to perform a full squat. For the first 10 answers I think the word "normal" is hard to agree on or quantify. Maybe functional is better? I don't think these are redundant at all. While pathology is considered to be a global term, I think it is helpful as a defining characteristic associated with a healthy ankle. Joint mechanics and arthrokinematics are the same thing. Joint mechanics and joint arthrokinematics. I prefer the term arthrokinematics and osteokinematics, or physiological and accessory movements. Physiological motion is basically ROM that can be measured with a goinometer. Acessory or arthrokinematics are glides, spins etc. Laxity and instability seem to be somewhat redundant. Functional deficits and functional assessment seem to be redundant. Many items are related, but not necessarily redundant. The list might be bioled down to pain with activity or rest, perceived or assessed functional limitations, perceived participation restrictions and normal joint stability and mechanics (which encompasses arthrokinematics - lots of redundancy
Describe the functional characteristics of someone with a healthy/normal/non-injured ankle.	There is redundancy in the first 5 items about ROM. I think of these as impairments as opposed to functional activities. Patient can ambulate stairs and patient can run may be included in "patient can perform all desired activities" if stair climbing and running were desired. Normal is a difficult term. Normal for whom? same with Full. I think you can be functionally capable without FULL range but need a functionally normal range of

motion. You could also say that you have pain with some things but it is not affecting you functionally, depending on how you define functional.

Again, some of these functions are not dependent only on ankle function.

"Full dorsiflexion" and "Normal inversion and eversion" are redundant with normal "AROM" and "PROM" The last item is redundant with "can perform all desired activities"

My thoughts are that if you use a lot of the functional measures above when talking about the ankle you may say someone does not have a normal ankle if they can not run, climb stairs, etc. when it is something else limiting them. I think the context of the question matters. Lots of the measures of ROM, balance, strength are not necessarily functional but mostly are required for someone to be functional. It is difficult to create a scale that is functional without including these components but they were included in the last question.

There are redundancies within range of motion and patient-reported function. I think normal range of motion captures the elements of the range of motion questions rather than specific directions. Full dorsiflexion is important, but one issue facing us now is the most appropriate way to measure it. I would moreso say that no range of motion asymmetries would be more appropriate to state than normal range of motion.

Full DF, inv, ev is part of full ROM

ROM statements are redundant. You list specific movements and then all movements. Proprioception, balance and neuromuscular control are overlapping concepts. The question is how would you test each separately...joint position sense, balance tests etc.

	"Giving way", "patient feels stable" and "ankle is stable" are redundant. "balance" and postural stability" seem redundant Absence of impairments, functional limitation and participation restriction might capture the lot Patient can perform all desired activities (including sport, work, and ADLs) - Answered no as this may not relate to ankle joint fucntion ONLY
Describe the functional characteristics or limitations of someone with an unhealthy/acutely injured ankle.	The range of motion options seemed redundant. Unable to cut and unable to change directions may be redundant. I think decreased is a difficult word as decreased compared to what? I ticked 'no' for somethings as I did not know what they were eg slow-downs, back peddle, plant. I think items about the same variable are redundant eg jump should only have one and unable to jump probably encompasses more deficits Many of these are similar, related, not necessarily redundant. ie range of motion is discussed and then specific motions, overlap with proprioception, blaance, coordination, stability (all are different but overlap and likely felt to be redundant by patient) Items 2-5 all deal with ROM as does "increased pronation/supination" and "increased ROM". "Decreased moblility" encompasses several more specific items like "cut", "slow downs", "back pedal", "jump", "run", "stairs", "Decreased balance" and "unable to balance" are redundant. The former is preferred. "Neuromuscular inhibition", "altered neuromuscular drive", and "reflex inhibition" are redundant" OK, my issue with this is do you have to have all on the list. I love it that

	decreased ROM and increased ROM as well as joint restriction and joint laxity are included. I agree! Could be all of these things but if someone has decreased confidence do they have an ankle injury? Loss of ROM covers all of the individual motions! There are overlapping items, but not necessarily redundancies on this list. Again, I think the range of motion overlap can be collapsed, but in this case, targeting specific range of motions rather than general may be appropriate. With regards to the patient reported measures, it may be more appropriate to state "difficulty doing..." rather than "unable to do...". Cost related to income/health care is not functional. Several of the above are related if not redundant. Some are completely opposite, but either may be present in an acute ankle injury. Neuromuscular inhibition, reflex inhibition and altered neuromuscular drive all pretty much the same thing other than altered could mean increase in muscle tone. Postural stability and balance very similar I wanted to note that I believe that many of items are characteristic of an unhealthy/acutely injured ankle. I only choose those items that I felt met the "functional criteria" in the question stem. Impairments (include neuromuscular control) function and participation - time for a facot analysis?
What is your definition of functional ankle instability?	Decreased muscle strength and decreased peroneal muscle strength are somewhat redundant. I am not familiar of the correct values for the FADI to indicate functional ankle instability. However, if these are the correct values - then I would have selected them.. The FAAM has similar values reported in this population - I am aware of literature for the FAAM and not for the FADI. Sorry for my confusing answer. If the FADI reference is encompassing the FAAM - then you can consider that I would have "agreed".

	Some redundancy as discussed in previous answer The deficits have to be related to their history of ankle instability. For example, someone could have a low FADI score but it is because they are recovering from an Achilles tendon injury. Have to be careful how this question is worded. FADI should be replaced with FAAM. Martin et al published their validation of the FAAM (condensed version of the FADI) in 2005. We have to stop using the FADI for research purposes because the FAAM has better psychometric properties. Functional ankle instability is the inability to perform desired functional activities due to real or perceived lack of ankle performance. This could be related to an unlimited variety of causes including but not limited to pain, swelling, strength, ROM issues limited or laxity, or perception, fear, confidence, etc. One of the issues with FAI for me when I was filling this out was that I don't feel I can dichotomize my responses because of the variations in FAI definitions and descriptions. Rather, for my disagree responses, I would say "...not necessarily". Perhaps these might be issues with those who report recurrent ankle sprains, but the overlap between MAI and FAI is often very large. This was a tough one!! :-) Functional instability is a clinical instability (instability found on clinical evaluation) that also causes a loss in normal function (ie causes the pt to seek help for the problem) FADI and CAIT likely capture FAI but in a simple sentence the key word is "function"
Describe the functional characteristics or limitations of	I think there is not a definite yes no for some of these but more a 'may be present' eg inhibited proprioception may be present but does not have to be (or at least has

someone with functional ankle instability; what are the functional costs or consequences?	not yet been demonstrated to be the case) Patients with chronic ankle instability (or FAI if you must call it that) are heterogenous. They are not all going to look the same. I agreed with most of the items here but that doesn't mean a patient must have each of these items to be classified as having FAI. It's a syndrome that consists of a common history of ankle sprain and a collection of symptoms. Not everyone will have the same symptoms. There is lots of redundancy on this list. 2 items about muslce endurance are redundant. "Laxity" and "ligament stability" are redundant. Fear, avoidance, apprehension and lack of confidence as well as other psychological issues are a result of functional limitations. Someone with functional ankle instability may not be able to perform activities at various levels especially if they do not learn to adapt and overcome the limitations. I think this is missing that some people may have functional ankle instability but return to full activity at a high level and do not have recurring problems. That does not mean the functional ankle instability is gone/cured! Again, with these responses, there are redundancies, but they capture the overlapping issues that are important to consolidate. With my disagree responses, what I would really like to say is "not necessarily..." Characteristics are difficult because they may be different pt to pt. Mainly though, instability is ligamentous, capsular, or joint insufficiency. Motion can be limited or excessive and not all characterisitcs apply to all patients Abnormal gait - Depends on how qunatified (Likely to observed altered kinematics and kinetics in lab). Chronic swelling, Synovial changes, Arthrokinematic changes, Ligament laxity, Degenerative changes - These are more characteristic of MI

Do you have a different title or name that you use for functional ankle instability?	Chronic ankle instability and functional ankle instability are interchangable. I prefer the terms recurrent ankle instability or recurrent ankle dysfunction. Chronic instability and functional instability is not synonymous (or shouldn't be). I think that CAI is best. Then there are two separate but linked terms (FI and MI). Each are characterized by different insufficiencies. I would be an advocate of the Hertel (2002) paradigm. I then use the definitions outlined in Delahunt et al (2010) to describe FI and MI.
Do you feel that there is a name more appropriate than functional ankle instability?	Dysfunction seems to best capture function and participation and permit the inclusion of ankle OA
Do you feel that functional and chronic ankle instability are the same thing?	I also think that chronic is not really the appropriate term in that it commonly refers to a long standing unresolved problem. While this could be the argument for the perception of instability, those with the condition experience repeated bouts of acute issues mainly. I think it is therefore important to refer to the condition not as chronic, but as recurrent. From an instability standpoint, it is a self-reported condition in which local elements of joint instability may or may not be present, but global activity limitations and participation restrictions typically are present. That is why I prefer the term Dysfunction rather than Instability. Both may not be fully descriptive. Is the chronically symptomatic ankle always unstable? Likewise, does instability always lead to symptoms? The dilemma is recurrent vs chronic (some can participate and get reinjured, others are permanently limited (chronic pain, swelling etc)

What is your definition of chronic ankle instability?	I think it is an umbrella term and all the 'agrees' can be part chronic AI but do not neccessarily define it. Any of these symptoms can be present in a patient with CAI, but they all don't have to be present for CAI to exist. Again, there are no items that are entirely redundant. Repeated, or permanent symptoms resulting in functional and participation loss
What are the functional characteristics of someone with chronic ankle instability; what are the functional costs or consequences?	Same as functionla but may include a wider range of characteristics. Some folks become hypomobile yet experience ankle dysfunction
Do you have a different title or name that you use to describe chronic instability?	I like dysfunction as an adjective
Do you feel that "persistent ankle instability" is an appropriate substitute for the title or name that you use (e.g. functional ankle instability, chronic ankle instability, ankle instability, etc.)?	Instability may not be the most appropriate term to use as the condition often manifests as a local issue initially, but translates to a global disability. Dysfunction May be a more appropriate term in my opinion. Persistent may be a better word than chronic, but it still may or may not be a functional instability.

	Persistent suggests that others are transient or temporary. At any given point in time the ankle is either stable or unstable, symptomatic or unsymptomatic, has full ROM or not, etc. Finding one term to classify all ankles with chronic problems is difficult unless you just say it's a dysfunctional ankle and then list specific impairments. I found the survey frustrating because any of the answers could be possible for any given patient. For research purposes, the key is to make subject populations as homgenous as possible. I don't think any of these terms do that. It's like saying someone has low back pain. There could be multiple combinations of underlying pathologies and impairments that lead to the production of low back pain. Research requires the identification of specific impairments that rule patients in or out. Clinical prediction rules are one way of doing this.
	Yes, persistent ankle instability can substitue for chronic ankle instability, but it cant replace functional or mechanical ankle instability
	Dysfunction over instability

Study 1, Round 3 Expert Written Comments

Question	Comments
Do you agree with this definition of a healthy ankle?	Many of the terms you use evoke the terminology the World Health Organization has adopted in its disablement model. They talk about disability (current terminology is participation), functional limitations (current terminology is activities), impairments (body functions and structures) and pathology (current terminology is health condition). You might want to reword to be in concert with

	the world health organization because you are close already. I believe pain is considered an impairment. Something like "A person with a healthy ankle, either through self-report or clinician examination, presents without pathology, impairments, functional limitations or disability relative to the ankle". That's using the old WHO terminology. The new definitions which were meant to be more inclusive of all health conditions may require alternative wording. The old categories made sense to me. I don't disagree with the definition but wanted to ask the question about why pain was listed specifically although it is considered an impairment. Philosophically, I agree that the most common impairment is pain, but it is none the less an impairment. Impairments of what? Would also put functional deficits last. Start w pain, impairments (deficits) etc then end w functional impairments I do not see the need to include pathology. Could a concerning pathology exist without impairment, functional loss or participation restriction? Also, for clarity, pain is an impairment.

Study 1, Round 4 Expert Written Comments

Question	Comments
This question addresses the assessment techniques listed below. The list is based on the panel's response to the original interview question "How do you	Is it analyze motion or "motion analysis"? Just wanted clarification.

assess functional characteristics of the ankle (for injured or healthy ankles)?" *I am now asking you to "agree" or "disagree" with each technique.	"Jog without pain" also "backpedal without pain", if the patient has pain but can still function vs pain interfering with the ability to function. So just can they jog or backpedal! Willingness to do activities, again are they unwilling because of the ankle? "for distance" with jumping or hopping, not necessisarily! Also not familiar with the 3 hop test or the hop and stop test. For balance testing I think single leg stance, star, etc. are done but we are talking "function." This stuff should be done before we are testing function and if it can not be done stop the test. Or if you try higher level function testing and they can not do it some of this stuff is to figure out why they can not function. I kept going back and re-reading the question toward the end. The reason I put disagree for most of the ROM, strength, palpation stuff is that based on the question I would have been beyond that part of the testing if I am looking for function. If you do every test you know you will get false positives, ie. Rhomberg is a brain injury test, not an ankle propriocetion test. Yes we need to document impairment with testing and then progress with testing again, but we are NOT using that stuff to determine function!!
	Lots of redundancy here.
	I answered this question as what we would actually do rather than what I think can be done. Happy to do again to put in what I think can be done. Some tests I ticked even though we do a variation of them eg gastrocs strength we actually do an endurance test, SF36 we use the AQoL which is an Australian equivalent. I still dont know what a back peddle is!!
	Lots of redundancy. Many of my disagrees are based upon that fact that they

require equipment note easily available (motion analysis system, EMG etc.). Rhomberg, tandem, SF-36 not specific enough tests. All other tests I could use in the right setting with the right patient, but I don't use them all, all the time.

Based on what has emerged, there were several redundancies associated with the assessments. Rather than choosing one and eliminating others, I selected all that might be related to one another. What I really like about this is that it seems as though laboratory-oriented, clinician-oriented, and patient-oriented measures have emerged as being relevant for assessment. This is a great sign that we're looking for context among the patient, clinician, and researcher! Great work Kelli!! Pat

I was a little confused on any of the strength tests or mobility/agility test measures. I think all tests are valid here if you are looking to see if the athlete has pain or not while performing them. I am assuming your are looking for performance on the strength/agility/mobility tests for pain free. I think they are all valid if you plug in the "phrase performed without pain".

As I'm sure you know, there is a lot of redundancy on the list but I imagine you'll be narrowing the items as we go. For clarifiction, the FAAM is the newer version of the FADI and the discontinuation of the FADI should be encouraged. Martin et al (2005) performed an item analysis of the FADI on nearly 1000 foot and ankle injured subjects. They eliminated a few items that did not provide unique information and the survey instrument with the retained items were re-named the FAAM. We should all be using the FAAM because there are items on the FADI that have been shown to be "invalid".

Lots of overlap the could be culled to something like self-reported function with or without standardized instruments (eg FADI) presentation (Ottawa ankle rules) pain and impairment measures plus physical exam of motion and stability functionassessment (walking, hopping, stairs, running cutting) particpation assessment (can they play their sport)

Study 1, Round 5 Expert Written Comments

Question	Comments
Do you agree or disagree with the following definition of functional ankle instability: "Functional ankle instability is a recurrent sense of giving way of the ankle."	I think there still exists much confusion about the use of the term functional ankle instability. 1) based on the contemporary models for explaining recurrent ankle instability, functional insufficiency is an aspect of the clinical phenomenon, but does not explain the phenomenon entirely. Recurrent ankle instability is multi-factorial and while it was originally described and defined by Freeman as FAI, the plethora of research about this phenomenon leads me to believe that we may need a more appropriate term to advance the awareness and understanding of this clinical phenomenon. I don't necessarily think that "chronic" is an appropriate term, but I do feel that recurrent may be more descriptive of the phenomenon rather than functional. There are functional deficits associated with the phenomenon, but the repeated episodes of giving way lead me to believe that it is a recurrent phenomenon with functional consequences. These functional consequences often manifest in other lower extremity phenomena as well (impaired proprioception, postural control, gait, performance, etc.) So the term functional is more so a description of an aspect of the phenomenon, but not the phenomenon itself. I hope that makes sense Kelli!!
	Maybe the best definition that could be offered
	I would add "or actual giving way" before "of the ankle." I would also remove recurrent because using recurrent makes it "Chronic Functional Ankle Instability." If you have an episode of functional ankle instability that resolves it is not chronic!!!
	This definition is too simplistic. At a minimum, the definition has to include reference to the lingering effects of an initial injury and current functional limitations because of their ankle. In my opinion, FAI absolutely has to be

	defined as a sequala to an initial lateral ankle sprain. How can there be instability if there is no initial injury to the lateral ligaments? What's unstable? I can't stress to you how important this is if you want your FAI definition to be accepted in the sports medicine community. Additionally, the functional limitations and characteristics that were agreed upon by the panel need to get into the definition of FAI; not as a whole list but as a concept that the patients cannot do the activities that they want to do. I agree that recurrent sense of giving way can be one aspect of functional AI but I do not agree that this is the only definition. What about recurrent sprain, or feeling of instability but not giving way? In the same way I do not agree that the functional characteristics should imply that all people have the characteristics but rather these are characteristics that MAY occur-not that you have to have them all.
Do you agree or disagree with the following definition of chronic ankle instability: "Chronic ankle instability is recurrent episodes of ankle instability and sprains."	The recurrent sprain is an important distinction from functional instability "Chronic ankle instability is RECURRENT episodes of functional ankle instability..."feeling of giving out or actual giving out." An individual with chronic ankle instability does not necessarily sprain because they have torn the ligaments previously. So I think you need to take out "and" and replace it with OR. There have to be functional limitations associated with the definition of CAI as well. I think CAI embraces all forms of instability which may have subsets of mechanical instability, functional instability etc. I do not feel it is restricted to

the above. UNLESS the definition of ankle instability is a broad one that encompases real and perceived instability and giving way.

You can't have recurrent episodes of ankle instability. If you do, it would be functional instability. Mechanical instability will always be there unless stabilized by arthritis or surgery. Neuromuscular control over mechanical instability that prevents giving way episodes in spite of mechanical instability is functional stability. Recurrent is a bad term. Recurrent mechanical instability doesn't happen. It is either there or not. Once there, it is always there until stabilized as mentioned above. Recurrent episodes can therefore only mean functional instability and giving way episodes which can come and go. I feel strongly that the proposed term is not helpful and suggests nothing more that functional instability therefore only confusing the matter. Chronic instability is a bad term. There is either compensated mechanical instability or uncompensated mechanical instability. Just like ACLs. Either the person can control the mechanical instability they have through neuromuscular control or they can't. This goes for an individual with a mechanically stable joint. They can either control there joint mechanics or they can't. Functional instability is the better term to describe ongoing instability.