FootFAST screening tool for overuse injury risk: Psychometric properties of its three visual assessments of function.

Author
Marian Louise Baxter

Supervisory Team
David McBride
Paul Hendrick
Daniel Cury Ribeiro
Stephan Milosavljevic

June 2014

1 Abstract

Lower limb overuse injuries are common in the military, and their causes are multifactorial. Both internal factors such as anthropometrics and body mass index, and external factors such as equipment use and weapons carriage, have been implicated in the development of such injuries because of their potential detriment to efficient gait biomechanics. Visual assessments of function have been used as a method of assessing such injurious biomechanics, and are commonplace in the clinical environment. Despite the prevalence of their use, the psychometric properties of such tests have not been investigated extensively. This is an issue, as a minimum level of reliability and validity needs to be reached to ensure practises are justified.

A recently developed visual assessment of function, FootFAST, had not been investigated in such a way. FootFAST was designed to be used with populations such as the military, and specifically aims to predict lower limb overuse injury risk. FootFAST comprises three visual assessments of function- hindfoot angle assessment, single leg stance and heel raise.

The standard intervention method utilised with FootFAST, the shoe-inserted foot orthotic, had not been rigorously investigated in the population of interest, the military. Thus, the effectiveness of orthoses, and the justification of their continued use was uncertain. In order to address the issues outlined, the current thesis aimed to answer five research questions, which were:

1. What is the current evidence of the clinical effectiveness of orthoses in reducing target injuries in the population of interest: the military?
2. What is the current evidence for psychometric properties of clinical assessment tools for injury risk identification in the lower limb?

3. What is the inter-rater reliability of the three tests within the FootFAST screening protocol?

4. What is the intra-rater reliability of the three tests within the FootFAST screening protocol?

5. What is the criterion based validity of the three tests within the FootFAST screening protocol?

In order to answer these questions, the thesis used a multi-method approach and is reported in a hybrid structure. To answer question one, a systematic review including meta-analysis of current literature was completed. It was found that current evidence did not support the continued use of orthoses as an en-masse intervention for the prevention of lower limb overuse injuries in the military. The second question was addressed by a narrative review of current literature which investigated inter-rater reliability, intra-rater reliability and/or validity of a visually-based, functional assessment. It was found that the popularity of visual assessments of function greatly outweighed current knowledge, and in some cases, the results question their continued use in a clinical environment. In order to address the final three research questions, a feasibility study was conducted which then informed the methodological choices and powering of the main research study. The results of the main study, which included eighteen assessors (comprised of sports medicine practitioners, physiotherapists, and New Zealand army medics) and eighteen participants, indicated that two out of the three tests within FootFAST were suitable for continued use in their current form. More research is needed in this area in order to address the issues identified within the thesis.

2 Acknowledgements

It is not the crossing of the finish line that gives you such satisfaction; it is every step you took in order to reach it.

I would firstly like to thank my supervisory team for all their hard work and effort in helping me with my thesis. David, my primary supervisor and source of endless military contacts has facilitated data collection to occur without delays. Stephan, my secondary supervisor and 'within-a-week' reviser of Chapters for the first two years of the thesis: I don't know if I would have ever finished in time if it wasn't for his support. Paul, always pushed me further and ensured that my work kept getting better. Finally Daniel, gave me more time than he should have, making sure I left no stone unturned, and providing up-to-date thesis knowledge, a HUGE thank you.

Next, I would like to extend a huge thank you to Charles Baycroft. I met Charlie while I was researching for my honours project, and since then he has provided me with on-going support throughout my research into lower limb overuse injury prevention. The protocol which is investigated within the thesis was his creation, and is a demonstration of his intuitive skill in injury risk identification. Charlie was always there with a sharp eye for detail, and an endless source of knowledge.

None of the work would have been possible had I not been awarded a scholarship from the Foot Science Foundation in Christchurch. Owning a house and needing to eat can present big barriers to someone who decides to spend the next three years earning no money and thinking too much about one thing. The Foundation has been nothing but perfect in terms of providing a scholarship: never asking for too much of my time, supporting all my decisions, encouraging me (and occasionally funding me) to present at conferences, and of course, reliably providing my scholarship on the 20th of each month!

Lastly, but by no means leastly, I need to thank my parents, for providing me with support (both emotional and financial) throughout my studies. It goes without saying that I wouldn’t be here if it wasn’t for them. Especially my dad, who has been behind-the-scenes providing his expert knowledge and opinion, and even took up the tedious task of reading the entire thesis prior to submission.

3 Contents

1 Abstract ii

2 Acknowledgements iv

3 Contents vi

5 Tables xvi

6 Figures xviii

7 Abbreviations xx

8 Preliminary comments 1

8.1 Introductory Chapter and Literature reviews 2

8.2 Trials: Psychometric testing 4

8.3 In conclusion 5

9 Chapter 1. Introduction 7

9.1 Overview of Chapter 7

9.2 Risk of injury and efficiency of motion 7

9.3 Prevention is better than the cure 9

9.4 The movement of interest 9

9.5 The injuries of interest 10

9.6 Analysis of gait to predict injuries 11

9.7 Justification for the use of visual assessments 12

9.8 Reaching conclusions when using a visual assessment 14

9.9 Limitations of visual techniques 14

9.10 FootFAST - A way forward? 15

9.11 Research questions 16

9.12 Research hypotheses 17

9.13 In the next Chapter 17

10 Chapter 2. Overuse injury aetiology, risk factors and injury prevention 18

10.1 Overview of Chapter 18
10.2 Mechanics of injury in the lower limb 18
10.2.1 Fatiguing movements 19
10.2.2 Excessive movements: Joint Instability 19
10.2.3 Poor shock absorption 20
10.2.4 Equipment use 21
10.3 A history of 'abnormal foot mechanics' and injury risk 22
10.3.1 Abnormal foot pronation 22
10.3.2 Other abnormal foot movements 23
10.4 The ankle joint and injury occurrence 24
10.4.1 Instability of the ankle 24
10.5 Decreased proprioceptive ability 25
10.6 Physical and anthropometric risk factors 28
10.6.1 Fitness level 29
10.6.2 Sex 29
10.6.3 Ethnicity 29
10.6.4 BMI 30
10.6.5 Arch height 30
10.6.6 Summary of risk factors 31
10.7 Injury occurrence in the military 31
10.7.1 Unique character of the military training environment 32
10.8 Military footwear and physical activity 33
10.8.1 Boot ergonomics 34
10.8.2 Standard issue footwear 36
10.8.3 Physical training 38
10.9 Final comments: injury aetiology 39
10.10 Overuse injury prevention through risk factor identification 40

10.11 Military training....41
10.11.1 Intervention strategies for injury prevention....42
10.11.2 Stretching programs....42
10.11.3 Physical training....43
10.11.4 Concurrent training....43
10.12 Summary on interventions....44
10.13 Limitations within the literature....44
10.14 In the following Chapter....45
11 Chapter 3. A review of functional screening tests....46
11.1 Overview of Chapter....46
11.2 Background....46
11.3 Search strategy....49
11.3.1 Inclusion criteria:....49
11.3.2 Exclusion criteria:....50
11.4 Results....51
11.5 Studies identified....51
1.1 Identification....51
1.2 Screening....51
1.3 Eligibility....51
1.4 Included....51
11.5.1 Star Excursion Balance Test (SEBT)....52
11.5.2 The multi-hop test....53
11.5.3 Hip Joint Angle Testing (HJAT)....54
11.5.4 The 4-hop test and the T-test....54
11.5.5 The Balance Error Scoring System (BESS)....55
11.5.6 Gillet test....56
11.5.7 Hip extension test....57

11.5.8 Hindfoot alignment assessment 57
11.5.9 Proprioceptive ability or balance control: The single leg balance test 59
11.5.10 Sport specific tests 60
11.6 Discussion 61
11.6.1 Limitations of included studies 62
11.7 Conclusions 63
11.8 Summary 65
11.9 In the next Chapter 66
12 Chapter 4. Orthoses as an injury prevention strategy for lower limb overuse injuries in the military: A systematic review and meta-analysis of effectiveness 67
12.1 Overview of Chapter 67
12.2 Introduction 67
12.2.1 Objectives 68
12.3 Methods 69
12.3.1 Eligibility criteria 69
12.3.2 Population: Types of participants 69
12.3.3 Intervention: Types of interventions 69
12.3.4 Comparison: Types of comparisons 70
12.3.5 Outcomes: Types of outcome measures 71
12.3.6 Information sources 71
12.3.7 Search 72
12.4 Data collection and analysis 72
12.4.1 Study selection 72
12.4.2 Data collection process 72
Data items 73
12.4.3 Risk of bias in individual studies 73
12.4.4 Summary measures 74

12.4.5 Synthesis of results 74
12.5 Results 75
12.5.1 Study selection 75
1.5 Identification 75
1.6 Screening 75
1.7 Eligibility 75
1.8 Included 75
12.5.2 Study characteristics 77
12.5.3 Risk of bias in included studies 77
12.5.4 Between assessor agreement 78
12.5.5 Meta-analysis and synthesis of results 78
12.6 Discussion 88
12.6.1 Summary of evidence 88
12.6.2 Comparisons with other reviews 89
12.6.3 Effectiveness 90
12.6.4 Clinical relevance 90
12.6.5 Strengths and limitations of this review 91
12.6.6 Implications for research 92
12.7 Conclusions 93
12.8 In the next Chapter 94
13 Chapter 5. The FootFAST protocol 95
13.1 Overview of Chapter 95
13.2 Introduction 95
13.3 The FootFAST protocol 97
13.3.1 Test 1 (Hindfoot angle) 97
13.3.2 Test 2 (Single leg stance) 98
13.3.3 Test 3 (Heel raise) 98

13.4 A justification of the three tests comprising FootFAST (1) 99
13.4.1 Test 1 (Hindfoot angle) 99
13.4.2 Test 2 (Single leg balance) 101
13.4.3 Test 3 (Heel raise) 106
13.5 Final comments 107
13.6 In the next Chapter 108
14 Chapter 6. What are the psychometric properties of the injury risk assessment tool: FootFAST? A feasibility study 109
14.1 Overview of Chapter 109
14.2 Introduction 110
14.3 Methods 110
14.3.1 Participants 110
14.3.2 Inclusion criteria 110
14.3.3 Exclusion criteria 111
14.3.4 Assessors 111
14.3.5 Inclusion criteria 111
14.3.6 Exclusion criteria 111
14.3.7 Equipment 111
14.3.8 Kinematic data 112
14.3.9 Procedure 112
14.3.10 Test 1 (Hindfoot angle) 114
14.3.11 Test 2 (Single leg stance) 115
14.3.12 Test 3 (Heel raise) 116
14.4 Data collection and analysis 116
14.4.1 Continuous and categorical scales 117
14.4.2 Inter-rater reliability 118
14.4.3 Intra-rater reliability 118

14.4.4 Criterion based validity 119
14.4.5 Feedback from assessors 121
14.5 Statistical Calculations 121
14.5.1 Inter-rater reliability 121
14.5.2 Intra-rater reliability 122
14.5.3 Criterion based validity 122
14.6 Results 122
14.6.1 Criterion based validity 126
14.6.2 Musculoskeletal clinicians 126
14.6.3 Feedback from assessors 127
14.7 Discussion 127
14.7.1 Main findings 127
14.7.2 Assessors 128
14.7.3 Participants 128
14.7.4 Feedback from assessors 129
14.8 Further implications for the main study 130
14.8.1 Assessor experience 130
14.8.2 Sample size estimation 131
14.9 Limitations 132
14.10 Conclusions 133
14.11 In the next Chapter 134
15 Chapter 7. What are the psychometric properties of the injury risk assessment tool: FootFAST? A repeated measures cohort study in a military population 135
15.1 Overview of Chapter 135
15.2 Methods 136
15.2.1 Participants 136
15.2.2 Inclusion criteria 136

15.2.3 Exclusion criteria 137
15.2.4 Assessors 137
15.2.5 Inclusion criteria 138
15.2.6 Exclusion criteria 138
15.2.7 Equipment 138
15.2.8 Procedure 139
15.2.9 Test 1 (Hindfoot angle) 139
15.2.10 Test 2 (Single leg stance) 140
15.2.11 Test 3 (Heel raise) 141
15.3 Data collection and analysis 141
15.3.1 Inter-rater reliability 142
15.3.2 Intra-rater reliability 142
15.3.3 Criterion based validity 142
15.4 Statistical calculations 142
15.4.1 Inter-rater reliability 142
15.4.2 Intra-rater reliability 142
15.4.3 Criterion based validity 144
15.5 Results 144
15.5.1 Inter-rater reliability 144
15.5.2 Intra-rater reliability 147
15.5.3 Criterion based validity 148
15.6 Discussion 148
15.6.1 Summary of key findings 149
15.6.2 Inter-rater reliability; effect of participants 150
15.6.3 Inter-rater reliability 151
15.6.4 Intra-rater reliability between assessor groups 159
15.6.5 Intra-rater reliability 159

15.6.6 Criterion based validity 163
15.6.7 Summary of psychometric findings 167
15.7 Clinical relevance 168
15.7.1 For practitioners 170
15.8 Strengths and limitations 171
15.8.1 Strengths 171
15.8.2 Limitations 172
15.9 Future research directions 175
15.9.1 Revision of Test 3 175
15.9.2 Participants 178
15.9.3 Assessor training 179
15.9.4 A critique of FootFAST 179
15.9.5 Implications and recommendations for clinical practise 180
15.10 In the next Chapter 181
16 Chapter 8. Thesis Summary 182
16.1 Research Question 1 182
16.2 Research Question 2 184
16.3 Research Questions 3-5 185
16.3.1 Reliability 185
16.3.2 Validity 187
16.3.3 Summary 188
16.4 Final comments 189
17 References 191
18 Chapter 6, Appendix 1 203
19 Chapter 6, Appendix 2 204
20 Chapter 6, Appendix 3 209
21 Chapter 6, Appendix 4 213

22 Chapter 7, Appendix 5 220
23 Chapter 7, Appendix 6 222
24 Chapter 7, Appendix 7 228
25 Chapter 7, Appendix 8 234
26 Appendix 9, Ethical approval 239

5 Tables

Table 8-1: List of published material which has been produced from the thesis. 6
Table 11-1: Collected measures of psychometric properties from clinical tests of included studies. ICC; inter class correlation coefficient, K; weighted kappa score. 52
Table 12-1: Numbered items of quality assessment as per the PEDRO scale (165). 73
Table 12-2: Characteristics of the eight studies included in this review 76
Table 12-3: Assessment of quality for each of the studies included in the systematic review. 78
Table 14-1: Participant characteristics (N=10). 123
Table 14-2: Assessor characteristics (N=6) all with a minimum of three years clinical experience. 123
Table 14-3: Fliess' weighted Kappa for inter-rater reliability of categorical data. 124
Table 14-4: Fliess' weighted Kappa for inter-rater reliability of continuous data. 124
Table 14-5: Cohen's weighted Kappa (CWK) calculations for intra-rater reliability among the musculoskeletal physiotherapists, and average results, shown are for each test individually. 125
Table 14-6: Cohen's weighted kappa (CWK) values for the three tests within FootFAST, calculating criterion based validity (comparison of assessor results with Silicon Coach™ (hindfoot angle) and Gait™ (Single leg stance and Heel raise) software results). 126
Table 14-7: Inter and intra-rater reliability, among only musculoskeletal clinicians, assessed using Fliess' linear weighted Kappa scores. 127
Table 15-1: Fliess' weighted Kappa for inter-rater reliability of results assigned categorically 144
Table 15-2: Fliess' weighted Kappa for inter-rater reliability of continuous results.....145
Table 15-3: Flies weighted Kappa scores for inter-rater reliability, by participant, for each test 146
Table 15-4: Comparison of intra-rater reliability results attained between assessor groups. 147
Table 15-5: Cohen's weighted Kappa calculations for intra-rater reliability: averaged intra-rater results for all assessors (N=18), for each test. 147
Table 15-6: Cohen's weighted kappa (CWK) calculations for the three tests within FootFAST, calculating criterion based validity (comparison of assessor results with Silicon Coach™ (Test 1) and Gait™ (Tests 2 and 3) software results) 148

Table 15-7: Comparison of current results with the results of previous research 168

6 Figures

Figure 1: Diagram to illustrate the potential sources of overuse injury risk, as discussed in the current Chapter.....39
Figure 2: Diagram showing number of retained studies during the search process (number of retrieved and retained studies shown in brackets). 51
Figure 3: Diagram showing number of retained studies during the search process. Number of retrieved and retained studies are shown in brackets.....75
Figure 4: Forest Plot for Femoral Stress Fracture, Data obtained from three of the eight included studies. Data represents incidence of femoral stress fractures between experimental (orthoses) and control (no orthoses) groups.....79
Figure 5: Forest Plot for Stress Fractures of the Foot. Data obtained from four of the eight included studies. Data represents incidence of stress fractures of the foot between experimental (orthoses) and control (no orthoses) groups.....80
Figure 6: Forest Plot of General Stress Fractures (includes any lower limb stress fracture, with the exceptions of the tibia or femur). Data obtained from three out of the eight included studies. Data represents incidence of stress fractures referred to as 'general' between orthoses and control groups.....81
Figure 7: Forest Plot of Overuse Injuries of the Foot. Data obtained from two out of eight included studies. Data represents incidence of overuse injuries of the foot between orthoses and control groups.....82
Figure 8: Forest Plot of Overuse Injuries of the Back. Data obtained from two out of eight included studies. Data represents incidence of overuse injuries of the back between orthoses and control groups.....83
Figure 9: Forest Plot of Stress Fractures of the Tibia. Data obtained from three out of eight included studies. Data represents incidence of stress fractures of the tibia between orthoses and control groups.....84
Figure 10: Forest Plot of General Overuse Injury of the Knee. Data obtained from three out of eight included studies. Data represents incidence of general overuse injury of the knee between orthoses and control groups.....85
Figure 11: Forest Plot of Medial Tibial Stress Syndrome. Data obtained from three out of eight included studies. Data represents incidence of medial tibial stress syndrome between orthoses and control groups. 86

Figure 12: Forest Plot of Sprains and Strains. Data obtained from two of the eight included studies. Data represents incidence of sprains and strains between orthoses and control groups. 87
Figure 13: Forest Plot of Other Overuse Injuries. Data obtained from three of the eight included studies. Data represents incidence of any other overuse injury (unspecified) between orthoses and control groups. 88
Figure 14: Configuration of the data collection environment.....113
Figure 15: Example of the images captured during Test 1. 115
Figure 16: A still frame from a single leg stance video captured during Test 2.....116
Figure 17: Illustration of the concepts of centre of gravity and centre of pressure. 120

7 Abbreviations

ACC: Accident Compensation Corporation

BESS: Balance Error Scoring System

BMI: Body Mass Index

CAD: Computer Aided Design

CI: Confidence Interval

COP: Centre Of Pressure

CWK: Cohen's Weighted Kappa

DFHS: Defence Force Health School

DFMC: Defence Force Medical Centre

DCR: Daniel Cury-Ribeiro

DVD: Digital Video Disk

FootFAST: Foot Functional ASsessmenT

FWK: Fliess' Weighted Kappa

GRADE: Grading of Recommendations, Assessment, Development and Evaluation

GRF: Ground Reaction Force

ICC: Intra-class Correlation Co-efficient

JPS: Joint Position Sense

K: Kappa

Lt Col: Lieutenant Colonel

MLB: Marian Louise Baxter

MTSS: Medial Tibial Stress Syndrome

NZDF: New Zealand Defence Force

NZ: New Zealand

PEDRO: Physiotherapy Evidence Database

RCT: Randomised Controlled Trial

ROM: Range Of Motion

SLS: Single Leg Stance

SMNZ: Sports Medicine New Zealand

US(A): United States (of America)

8 Preliminary comments

This thesis reports a research programme focused on the development of FootFAST (FOOT Functional ASsessmenT), a simple screening tool based upon visual assessment. It is designed for use in the military to identify those at risk of lower limb overuse injury, and is based on the prescription of an intervention as an injury prevention strategy. The tool comprises three simple tests – hindfoot alignment, single legged stance, and heel raise.

Previous research completed with the New Zealand army in 2009 indicated that the FootFAST tool could potentially identify those at risk of overuse injury (1). Coupled with an orthoses intervention, the prevalence of defined injuries were also successfully reduced (1). The intervention supplied to those in the high risk group as part of this study was a self-moulding orthoses. The observed outcome was statistically significant reductions in the occurrence of ten predefined overuse injuries of the lower limb when compared to a parallel control group. Given such encouraging findings from this study, further investigations utilising the FootFAST protocol were considered to be justified.

The direction taken for the work reported in this thesis was to focus on the psychometric properties of the FootFAST tool. Although confidence in the clinical utility of FootFAST was high, reinforced by the encouraging findings from the reported study, data on the reliability and validity of the protocol were lacking. The protocol was relatively novel, having only recently been developed, and had very limited use – only the original developer (Dr Charles Baycroft, The Foot Science Foundation) and the author (through work on an honours project) had experience in applying the test. Thus, despite the encouraging results which suggested that the test validity was high, there was nothing to suggest that clinicians could use the test appropriately as this had never been investigated.

Given these issues, there was a clear direction for the thesis: an investigation into the psychometric properties of FootFAST. Test validity had already been supported by the previous investigation (1). The remaining properties – criterion based validity, inter-rater reliability, and intra-rater reliability, were assessed here to the highest possible methodological rigor.

8.1 Introductory Chapter and Literature reviews

Chapter 1 introduces the concept of overuse injuries and their prevention, and describes the concepts and components of the FootFAST tool; it concludes by highlighting the research aims for the thesis. Chapter 2 leads on from this, providing a background on overuse injuries and injury prevention in the lower limb, discussing overuse injury aetiology, and examines research which focusses on environmental and internal risks. It includes a focussed discussion on the population of interest, the military. The information collated in this Chapter was presented at the "*Southern Physiotherapy Symposium"*(2). A review of current literature regarding psychometric properties of functional assessments of overuse injury risk was then completed as background to the current research, presented in Chapter 3. Given the scarcity and lack of depth of studies concerning the latter topic, guidelines suggested it would be more appropriate to be written as a narrative review (3). Despite such paucity of data, findings were generally encouraging – almost all findings were in favour of the continued use of the tests. The information from this Chapter was presented at the "*Australian Conference of Science and Medicine in Sport"* (4), and has been accepted for publication in *Physical Therapy Reviews* (pending).

The second review completed for the thesis was a systematic review of the current evidence for orthoses interventions in the reduction of injury rates in the military, and is reported in Chapter 4. This review was not only undertaken to establish the current evidence

base for orthoses management in this population, but also to assess the intervention approaches undertaken, and the current quality of work within the field.

The quality and availability of studies identified as part of this review allowed only limited conclusions to be drawn. This was especially true when considering the allocation process utilised. In all of the studies identified, orthoses were prescribed based on a randomised allocation process, and not as targeted interventions based on risk assessment. Thus, the findings from the systematic review were not entirely generalizable to targeted interventions, which are 'prescribed' based upon screening tools such as the FootFAST protocol. In summary, it was concluded that although there was no evidence that intervention caused harm, there was only limited evidence of their preventive ability when used as a non-targeted intervention: only two out of ten studies showed substantial reduction. This Chapter has been published in *Physical Therapy Reviews* (5).

Although this was an important finding, it was considered to be limited in terms of relevance to the development of FootFAST and its approach to orthoses intervention. The prescribed orthoses (as utilised in the previous study (1)) was the current intervention of choice in the NZ army; these orthoses were practical in their efficiency of issue, and were supplied at a relatively affordable rate. In essence however, FootFAST could be used with any number of interventions based on the needs and requirements of the group or patient in question. Although the knowledge created during the systematic review was relevant, it is important to note that FootFAST was never solely designed to be utilised with orthoses interventions specifically.

Chapter 5 contains a detailed discussion of the FootFAST protocol, and the justification of the three visual assessments of function which comprise it: Hindfoot angle (Test 1), single leg stance (Test 2), and heel raise (Test 3). This information sets the background and provides a lead-in to the following Chapters: the feasibility study and the main study. FootFAST and

its justification were presented at the "*Australasian Podiatric Societies' Annual Conference"*(6).

8.2 Trials: Psychometric testing

The two main empirical studies reported in this thesis were trials of the psychometric properties of the FootFAST screening tool. The first of these studies (Chapter 6) was a feasibility trial of the proposed methodology, which also provided data to inform a power analysis for a full study of the psychometric properties of the FootFAST tool. The study provided feedback from the participants, which proved to be invaluable in terms of further developing the methods for the full study. Perhaps the most important issue identified through the feasibility study was the sometimes ambiguous wording of the instruction sheets for assessors. This would have caused difficulty had it not been amended prior to starting the main study. The results from this study were presented at the "*Australian Conference of Science and Medicine in Sport"* (7). A poster containing the results from this study was presented at the "*Dunedin School of Medicine Colloquium"* (8), and a manuscript for the study is currently under review with the *Journal of Physiotherapy*.

The main study (Chapter 7) required eighteen participants and eighteen assessors. This was determined previously using the results from the feasibility study, in order to reach statistical power. This calculation was effective, as no confidence intervals obtained during the statistical analysis were bigger than 0.20. Assessors were provided with videos and pictures instead of viewing patients in real-time. This meant that bias was minimised, and blinding and randomisation were possible. The results indicated that two out of the three tests within FootFAST were suitable for continued use in their current form, and the third requires revision in order to improve the level of criterion based validity observed.

8.3 In conclusion

If the current thesis' findings could be summarised in one paragraph, it would read as follows. There is not enough research which considers the psychometric properties of visual assessments of function, and research which has been completed to date shows varying results. The use of orthoses as an intervention for patients at high risk of overuse injury is not supported by current evidence, but there is no evidence that their use causes harm. However, a major weakness in previous studies and current approaches is the lack of adequate screening or prescription of orthoses interventions. The FootFAST protocol is designed as a screening tool to identify those at risk for certain types of injury, and thus as candidates suitable for an orthoses intervention. The current thesis found that the tool has clinically acceptable levels of inter and intra-rater reliability, and thus its use is justified in these respects. Of the three tests which comprise the tool, Tests 1 and 2 had acceptable levels of criterion based validity; however, Test 3 fell short of the standard. It is advised that Test 3 undergoes development prior to its continued use.

Table 8-1: List of published material which has been produced from the thesis.

Chapter	Paper title	Authors	Contribution of candidate	Journal or Conference	Status
1	Lower limb and lower back overuse injury prevention: Evidence from a specialised population.	Baxter, M L McBride, D I Milosavljevic, S Hendrick, P	First author, presenting author	Southern Physiotherapy Symposium	Presented in Queenstown, New Zealand 2011
3	Psychometric properties of lower limb and lower back injury risk assessment: A narrative review	Baxter, M L McBride, D I Cury-Ribeiro, D, Hendrick, P	First author, presenting author	Australian Conference of Science and Medicine in Sport	Presented in Phuket, Thailand 2013
3	Psychometric properties of lower limb and lower back injury risk assessment: A narrative review	Baxter, M L McBride, D I Cury-Ribeiro, D, Hendrick, P	First author	Physical Therapy Reviews	In Press
4	Do orthotics work as an injury prevention strategy for the military?	Baxter, M L Milosavljevic, S Cury-Ribeiro, D	First author	Physical Therapy Reviews	Published 17(4), 2012
5	A background and justification of the novel injury risk assessment protocol: FootFAST	Baxter, M L McBride, D I Milosavljevic, S Hendrick, P	First author, presenting author	Australasian Podiatric Societies Annual Conference	Presented Sydney, Australia, 2012
6	Are visual assessments made by clinicians reliable and valid?	Baxter, M L McBride, D I Milosavljevic, S Hendrick, P	First author, presenting author	Dunedin School of Medicine Colloquium	Presented Dunedin, New Zealand, 2012

9 Chapter 1. Introduction

9.1 Overview of Chapter

This Chapter provides the background and context for the work described in the current thesis. It starts with an overview of motion and musculoskeletal overuse injury risk, focussing on those most relevant to the current thesis. It concludes with a brief description of the FootFAST screening tool (continued in Chapter 5), which is the central focus of the current work. The overall topic of this thesis is prevention of overuse injuries of the lower limb, referred to as the target injuries. For the purposes of this thesis, an overuse injury is defined as: '*an injury sustained as a result of excessive or repetitive micro traumas'* (9). This definition was chosen as it aligns with a biomechanical and physiotherapy-based framework.

9.2 Risk of injury and efficiency of motion

In the military, the observation of movement as a means of identifying those at risk of injury dates back to the post World War II period (10). Such observations were part of 'recruit medicals' and 'routine' medical checks, where functional movement analysis was performed by a practitioner, and an 'injury risk' calculated (10). Such movement screening was based on the theory that visual recognition of a 'graceful efficiency of no wasted motion' is intrinsically recognised as being the optimal movement pattern (11). Based on this theory, it can be argued that inefficient movement patterns can also be identified through such visual analysis (11).

Inefficiency of movement and injury risk, from a biomechanical perspective, are considered to be inherently linked. According to this perspective, if a person is performing a task with an observably inefficient or poor technique, they would be producing unnecessary movements. Any 'additional' movements are likely to use more energy than necessary,

thereby causing unnecessary musculoskeletal stress and fatigue (12). The muscle groups involved in these movements then transmit stress vectors to adjacent muscle groups. Such abnormal forces over a prolonged time frame may plausibly be implicated in the development of chronic injury. The identification of such inefficient movement patterns is traditionally performed by a physiotherapist (also termed a physical therapist). (13)

Traditionally, the physiotherapist has been a highly skilled 'qualitative' (or visual) assessor of functional movements, such visual assessments were then used as part of clinical reasoning processes to identify a diagnosis, and – in turn – provision of appropriate interventions and treatments for functional injuries (including overuse injuries) and disabilities. More recently, the visual assessments made by the physiotherapist have been influenced by, and in some cases replaced with, the use of quantifiable biomechanical analysis.

Compared to the long history of qualitative assessment within the clinical setting (14), quantitative assessment of clinical disorders within the discipline of biomechanics is still in many respects a novel research area. Relatively recent developments of high speed digital video cameras, specialised equipment (such as the RS Scan^{TM} foot scanning device), and software packages (such as Silicon Coach^{TM}) have been the main modes of advancement in technology-based movement analysis (15-17). Although technology-based movement analysis began life within the sporting field (18), it has recently found translational utility and use within the field of clinical physiotherapy (19). The reason for this was the recognition of the potential for these instruments to provide more reliable and valid clinical measurements, reducing the observational variability associated with both visual assessments of movement and educated estimates of joint force parameters (19, 20).

9.3 Prevention is better than the cure

A key role of the physiotherapist and other injury management clinicians is the provision of effective management interventions for musculoskeletal conditions, including overuse injuries (21). Interventions are classified as either secondary, made following injury or dysfunction; or, in preference, primary in order to prevent injury or dysfunction. If we are to consider the negative consequences of sustaining an injury: cost to the health care system, time off work, pain and inconvenience, and the increased likelihood of future injury, then it seems logical that if primary prevention is feasible, it is a much better option than treatment (1, 12, 22). In order to prevent musculoskeletal overuse injury, individuals are screened for injury 'risk factors', classically identified biomechanically through a functional screening assessment. The main difficulty in such assessment lies in standardisation: there are over one hundred documented functional screening tests outlined in the clinical physiotherapy literature (14, 23, 24).

9.4 The movement of interest

Walking, as the fundamental bipedal gait, is universal to almost every population group (with the possible exception of a small group of those with disabilities, for whom movement restrictions do not allow normal ambulation), and is a common focus for functional assessment in the clinical area (25, 26). The human skeleton is effectively designed for efficient bipedal locomotion (27). In order to investigate walking using biomechanical analysis, parameters and terminologies which will be used within the current study must first be outlined.

The most basic unit of bipedalism is the gait cycle (28). For the purpose of this thesis, a gait cycle is defined as: '*the rhythmic alternating movements of the two lower limbs, with the goal of moving the body forwards'* (28). In effect, this means the activities which occur between heel strike of one limb, and the sequential heel strike of the same limb. A gait cycle

for walking has four phases: touchdown, stance, toe off, and swing (28). ‘Touchdown’ refers to the very brief period of time when the heel initially strikes the ground. ‘Toe off’ refers to the instance when the toes lift off the ground, the last moment of foot-ground contact. ‘Stance’ is a phase which begins when the heel of the forwards limb makes contact with the ground and ends when the toe of the same limb leaves the ground. The ‘swing’ phase begins when the foot is no longer in contact with the ground (toe off), which is followed by a rapid hip flexion bringing the foot past the torso, then a slowing of the foot in preparation for the next touch down. The phase finishes at the sequential touch down. (28-30)

A singular gait cycle refers to a period of one complete movement from touch down to touch down on the same foot. Various points within a gait cycle are of particular interest to the movement analyst, details of such will be discussed further below. It is because of the universal applicability of walking, its relevance to overuse injuries, and the high saliency of walking and walking-related assessment performed by the practitioner, that such a walking-related assessment was chosen as the topic of this research.

9.5 The injuries of interest

As indicated above, this research focuses on overuse injuries of the lower limb, which are referred to as the target injuries. These were chosen as they are the injuries that are most often associated with gait and other forms of bipedal movement (31-33). Observation of gait to identify ‘abnormality’ is likely to be one of the more common visual assessments made by the physiotherapist when investigating chronic injuries of the lower limb. That is not to suggest that aberrations in (or a suboptimal) gait pattern are the only causes of such injury, but these are recognised as a common contributing factor (1, 34).

The target injuries of interest here are the same as those identified by Davidson (35) and Baxter (1): stress fractures of the lower limb, foot, pelvis or lumbar spine, chronic pain or discomfort in the lower limb, foot, pelvis or lumbar spine, general overuse injuries of the

lumbar spine, foot, pelvis or lower limb, and plantar fasciitis. Some specific overuse injuries have been excluded, for example any fascial problems (not including plantar fasciitis) with the lower limb, the rationale being that there is evidence that hereditary factors contribute to the development of these problems (36, 37). Given such differences in clinical aetiology, such injuries have been excluded, while those which are more likely to be environmental are included.

9.6 Analysis of gait to predict injuries

Traditional observational analysis involves the client or patient being asked to walk along a predefined walk-way or over a given distance whilst being observed from a distance (14). Biomechanics of walking can now be interpreted quantitatively, allowing the ability to capture (for example) force vectors, balance ability, and joint positions (19, 38). These factors can then be used in a model which predicts the risk of injury in a manner consistent with an objective, scientific (i.e. validated) approach (39, 40).

These quantitative measurements, known as gait kinematics, cannot be precisely derived from visual observations alone. Such kinematic knowledge has provided a detailed insight into many areas of the gait cycle: the sinusoidal path of the centre of mass, the peaks and troughs of flexion and extension in the knee during a gait cycle, the differences when walking barefoot compared to shod and the associated computer modelling based on such kinematics (41).

Advancements such as these are largely attributed to the advancements in technology. These technological advancements began with the development of computer assisted design (CAD) programs such as Silicon Coach^{TM} (SiliconCOACH, New Zealand), which incorporated the use of two dimensional (2D) and three dimensional (3D) camera systems. Silicon Coach^{TM} and other CAD programs gave practitioners the ability to accurately measure, track and quantify joint angles (19, 38).

More recently, the development of systems such as the RS Scan™ (RS Scan International, Belgium), provide a much more detailed description of certain kinematics during gait; to the extent that gait related injury risk can be calculated directly from their outputs (17). Previously, gait kinematics could only be estimated through visual observation.

Despite these advances, apart from specialised clinical units (e.g. teaching hospitals), it is still common practice to use visual gait assessment in the vast majority of clinical settings. This is because of two reasons: the costs involved in purchasing and maintaining such technology, and the level of observer training required. In addition to this, the intuitive analysis inherent to non-invasive visual assessments are more practical in everyday physiotherapy practise, for the sports medic, or for any other practitioner performing such analysis (42).

9.7 Justification for the use of visual assessments

In contrast to the quantitative, biomechanically-based modelling approach, an argument is that visual assessments are more appropriate for routine use, as they do not assume that there is some fundamental 'ideal model' for walking (42). It has long been recognised that no two people will walk or run in the same way, as there is an infinite amount of variability possible in the human bipedal system: gait-based identification of individuals is used in many areas because of this (43). Based on this argument, the idea that there is an ideal model for bipedalism is counterintuitive. Anthropometric differences, turn over speeds (stride rate), stride lengths, arm swing, and endurance are some of the many factors which dictate gait pattern (11). This then presents a problem for computer based biomechanical analysis; the nature of a human or biological system means that there will never be a singular ideal movement form or comparison model (11, 43), and approximations are as named: an approximation. A patient-specific three dimensional (3D) gait analysis using markers can overcome this issue, but then there is the added issue of correct marker placement, which is

known to be somewhat problematic within a clinical setting (44). Furthermore, 3D gait analysis still entails the generalised issues faced with technology-based analysis, as outlined in points two, four and five below.

The final point to raise here is that most actions and movements during gait are considered to be highly situationally specific (42). This is important as gait analysis equipment, specifically those which include plantar pressure measurements, typically require the individual to be barefoot. Although barefoot movements are an accepted baseline state for reference (45), the biomechanical analysis made whilst barefoot is not likely to be the same as when wearing footwear (46). A visual assessment has the advantage in this case, as there is no restriction placed on footwear or equipment choices; for example, rugby boots, track running shoes and other sprig-based footwear cannot be incorporated in investigations where plantar pressure mats are utilised, which potentially gives it a higher level of ecological validity (47).

To summarise, the primary arguments for using visual assessments of function instead of technology or software-based assessments in routine clinical practise are as follows.

1. The individual variability of motion, which at this stage cannot be interpreted by motion analysis technology,
2. A lack of knowledge or clear consensus of what is important and/or relevant to measure quantitatively,
3. A lack of ecological validity (e.g. footwear use) in biomechanical assessment,
4. No general consensus on normative values for the various parameters of interest,
5. Resource implications, including capital and running costs, and the need for highly skilled observers.

Based upon the discussion above, visual assessment is clearly a more viable means of assessing gait in the clinical setting.

9.8 Reaching conclusions when using a visual assessment

The primary goal of visual assessment when screening for musculoskeletal injury prevention is identification of the key repetitive stresses potentially responsible for future injury; for example, ligamentous laxities (48, 49). Repetitive stresses represent a key element of such screening, as clinical analysis of injury aetiology has estimated that 75% of lower limb injuries are from overuse-repetition of the same constant movement (45). When an injury does occur, it is suggested here that these movements are, for the individual, anatomically abnormal and thus pathophysiological. For example, it may be observed through a clinically based visual observation of gait that the hip structure tracks considerably more to the left and/or right during walking. This might be due to capsular laxity and over-reliance on the hip dynamic stabilisers (11). Using such a visually-based assessment procedure, a specific targeted intervention for this individual can be identified. It is postulated here that this process might not be possible, or would take longer if it were completed using, for example, footprint analysis software or dynamic movement analysis software such as Silicon Coach (SiliconCOACH Ltd, New Zealand) or Kinovea (www.kinovea.org).

9.9 Limitations of visual techniques

The arguments above illustrate the comparative simplicity and efficiency of using visual analysis to assess risk of musculoskeletal overuse injuries. However, visual analysis is not without its limitations. For a start, there are many subtle aspects of gait which cannot be clinically discriminated by the simple diagnostic algorithm of inspection. As discussed above, a footprint system, when used solely, cannot accurately identify the motions of the lower limb; equally, the eye cannot see what is occurring in the foot-ground interface. Ground reaction forces (an expression of the forces exerted on the foot during contact), centre of pressure, velocity of force vectors, and percentage body weight are all examples of

parameters which may be relevant to the risk of overuse injury; however, none of these can be assessed without the use of a quantitative analysis-based system. (11, 45, 50)

As there is no formalised reference point for the 'ideal form' during gait, many researchers have indicated that the conclusions made are entirely subjective and based on the opinion and experience of the clinician (51, 52). This naturally leads to variability in the results of such assessments. In contrast, calibrated software based systems have negligible amounts of variance (53, 54). Such variability is of concern for a number of reasons, including that it potentially limits the ability to standardise and replicate visual assessments, therefore limiting the ability to do collaborative research (55), and reducing the credibility of the use of visual analysis and thus confidence in the reliability of this approach.

The greatest issue which arises from the visual analysis of movement is that the reliability and validity of such analysis is questionable, as there is a lack of test standardisation. Taking the example of Test 2 from the FootFAST tool (i.e. single leg stance), there are several test protocols documented in the literature (56-59). In this situation the use of such analysis to inform subsequent identification of the appropriate intervention - and subsequent monitoring of applied interventions - are also fraught with difficulty (60). If it remains the case that traditional, visual based assessments of injury risk are the main tool within the physiotherapy clinic, then it is essential that the reliability and validity of these measures are investigated, if simply because the limitations of its use must be known.

9.10 FootFAST - A way forward?

As the visually-based functional assessment of overuse injury risk remains a clinically important tool, techniques or assessments utilised need to be validated. The protocol investigated here, named FootFAST (**FOOT F**unctional **AS**sessmen**T**), was recently developed and trialled within the New Zealand army. Results indicated that the tool had a

strong predictive ability, and eight out of ten overuse injuries were significantly reduced based on the protocol and intervention supplied (1).

The New Zealand military provides a unique milieu in which to assess the potential for injury prevention. Recent assessments indicated that at any time 10% of personnel are not able to attend work because of the overuse injuries (as outlined above) related to gait (1, 35). Many reasons have been postulated as to why those serving in the military experience such a high prevalence of overuse injuries, a problem shared by the US, British, Danish forces and others (61-63). This has included the high level of physical demands placed on the soldier, and the use of sub-optimal equipment (1, 35, 64-66). Regardless of the causes, it is clear that this unique group requires specialist attention in order to address the issue of overuse injuries among them. Because of this, the military will be the population of interest for the current study.

9.11 Research questions

This thesis aimed to address the following research questions:

1. What is the current evidence of clinical effectiveness of orthoses in reducing target injuries in the population of interest: the military?
2. What is the current evidence for psychometric properties of clinical assessment tools for injury risk identification?
3. What is the inter-rater reliability of the three tests within the FootFAST screening protocol?
4. What is the intra-rater reliability of the three tests within the FootFAST screening protocol?
5. What is the criterion based validity of the three tests within the FootFAST screening protocol?

9.12 Research hypotheses

The research had several hypotheses, which correspond to the five research questions.

1. It was hypothesised that the current literature would provide equivocal evidence for the use of orthoses as an injury prevention strategy for the military.
2. It was hypothesised that there would be a lack of comprehensive information regarding the psychometric properties of clinically based, functional assessment tools for identifying risk of overuse injury. Those which have been tested for reliability and validity were hypothesised to have moderate to good levels of intra- and inter-rater reliability and poor levels of construct and test validity.
3. The inter-rater reliability of the three tests within FootFAST were hypothesised to have an overall good level of agreement, based on the inter-rater reliability observed from other clinical tests (67).
4. The intra-rater reliability of the three tests within FootFAST were hypothesised to have an overall moderate level of agreement, based on the intra-rater reliability observed from other clinical tests (67).
5. The criterion based validity of the three tests within FootFAST were hypothesised to be poor to moderate. This was based on the limited available background evidence; the only available base for predicting this comparison was the other functional assessments that had been compared with a gold standard, which were known to not produce good results (1, 68).

9.13 In the next Chapter

The following Chapter discusses the aetiology of overuse injury from a biomechanical framework. It discusses both the current and historical perspectives on what constitutes suboptimal motion and outlines how these movements are implicated in the development of overuse injuries. The following Chapter is important in setting the background for the rest of the thesis.

10 Chapter 2. Overuse injury aetiology, risk factors and injury prevention

10.1 Overview of Chapter

This Chapter provides an introduction to overuse injury aetiology from a biomechanics perspective. The Chapter begins with a discussion of the mechanics of injury. It then discusses both current and historical theories of foot motion as related to injury occurrence. The Chapter then presents a discussion of overuse injury prevention through identification of risk factors. Various risk factors are identified and examined. Following this, the injury risk factors which are unique to the population of interest for this thesis, the military, are discussed. The Chapter then explores current research available on overuse injury prevention among the military. The Chapter concludes by briefly identifying specific limitations of current literature in this area.

10.2 Mechanics of injury in the lower limb

In normal functional circumstances, the foot and lower limb move in an efficient manner to produce movement (27). This includes any form of bipedal gait, jumping, hopping and landing motions, turning, and pivoting; all of these movements can be collectively referred to as locomotion. In demanding and functionally limiting circumstances however, the efficiency of locomotion can be compromised, and lead to injury; for example, when using sporting equipment or when in military training (69, 70). Demanding or functionally limiting circumstances can occur because of: biomechanical inefficiencies and/or ineffectiveness (for the purposes of the discussion, both inefficiencies and ineffectiveness will be collectively referred to as fatiguing movements), excessive motion patterning and poor shock absorption, and – especially in the military – inappropriate or poorly fitting equipment (64, 70). All of these functionally limiting circumstances will be discussed in this section.

10.2.1 Fatiguing movements

Excessive movements can be defined as those which are unnecessary to produce the desired locomotive outcome, which is very apparent in those with gait pathologies (71). In essence, such movements result in increased energy consumption which can then lead to premature fatigue (35, 71). Fatigue has been identified as a significant risk factor for overuse injury (35). However, in order to assess this, fatigue must be measurable. Fatigue of a structure – both the person or the anatomical structure – is evident and measureable in many ways, for example, Stolwijk (72) investigated the changes in lower limb mechanics and ground reaction forces following fatigue. It was found that when the lower limb becomes fatigued following moderate exercise, there is a significant shift in plantar pressure towards the rear of the foot (72). In these circumstances, the toes are unloaded and the metatarsals and heel are loaded over time; the centre of mass is then significantly displaced towards the rear of the foot (72). Evidence for such an altered loading pattern could explain the increased risk of overuse injury when the body is fatigued (35), and the need to prevent or prolong time to fatigue is therefore important for injury prevention.

10.2.2 Excessive movements: Joint Instability

Fundamentally, joint instability can be defined as either functional or structural (73, 74). Structural instability refers to some deficit in the structures supporting the joints, and can result in low levels of bony congruence (74). For example, increased excursion of the talus at the talo-crural joint and increased talar tilt are indicative of a structural instability (73), and the term is commonly used interchangeably with the term laxity. Structural instability, or laxity, of the joint is commonly seen in individuals with ligamentous deficiencies (74). Ligamentous deficiencies contribute to an increased risk of overuse injuries, particularly in cases were the deficiency is such that the joint cannot be adequately stabilised (75). Sprains and strains are likely to occur in these situations, and in order to facilitate stability the joint

must be maintained by the active contractile structures instead. Such demand will lead in turn to premature fatigue which, as indicated, is a potential precursor of overuse injury.

Functional instability refers to a deficit in the active, contractile structures supporting the joint which are under voluntary control (76). Specifically, instability arises when such structures allow movement to occur beyond the normal physiological range of motion (76). Such a deficit is typically seen with those who have weak peroneal muscle strength about the ankle joint (73), one example of a lower limb injury area of interest in the current thesis. The consequences of functional instability, as related to overuse injury risk, have been well described in the literature. Using the same example of the ankle joint, Dayakidis (76) observed that those with functional instabilities about the ankle joint had altered ground reaction forces during the initiation of cutting movements, which put them at increased risk of both acute and overuse injury.

Finally, it is important to note that structural and functional instability are not necessarily associated with each other (73); for example, Kerkhoffs (75) identified no difference in ankle joint *laxity* between groups of injured versus non injured players who had *functional* ankle joint instabilities.

10.2.3 Poor shock absorption

Another example of a potentially injurious movement is 'ineffective' or poor shock absorption through the restriction of force transmission; in order to efficiently displace forces encountered during locomotion, a kinetic chain of displacement must be functioning (77). The transmission of forces encountered during locomotion occurs along the 'kinetic chain', beginning at the foot and transferring throughout the body along the linked structures. The chain is in itself a structure whereby any force which acts upon one part will affect all other parts (77). Restriction of the chain in some way will ultimately affect all articulated

structures: for example, walking or running without swinging the arms leads to an excess of twisting from the hips in an effort to displace the forces (70).

Essentially, if shock absorption is poor or inadequate the structure will fail to efficiently dissipate the force imposed. If sufficient magnitude of force is applied to one point on a structure and shock absorption is poor, that specific point will experience the most compression and wear, therefore causing wear at uneven rates. Consequently, the location of the overuse injury is often the location where most wear is occurring. (12, 78)

The shock absorption required for certain locomotive activities makes them potentially more damaging than others. Based on a dose-response relationship, when the impact loading is maximal, i.e. during weight carriage or running, the injury risk is maximal. Anecdotally, many practitioners identify running as being the most damaging. The high impact shocks involved in running (1.5 to 5 times body weight (79)) have been correlated with injury statistics: it was reported that up to 70% of runners sustain an overuse injury every year (43-48% of which are injuries of the knee) (79).

Based on the topics discussed, the most injurious movement is likely to be carrying weight while running with restricted arm motion, and frequent, rapid changes of direction-tasks often included in the training and fitness testing of military personnel.

10.2.4 Equipment use

Potentially injurious circumstances can occur because of external impositions (69, 70). If the normal or natural movements of the lower limb and lumbar spine are restricted or compromised in some way, this may lead to changes in movement patterns, and in turn to inefficient movement or function (64, 70). For example, the wearing of a military boot, or arguably any footwear when compared to barefoot, causes angles of the foot and ankle joints to change: the ankle becoming more fixed due to the boot enclosing it (64, 70). Similarly, if the wrong size or incorrect fit of footwear is worn, termed ‘inappropriate equipment’, this

will restrict the normal splaying and bending of the joints of the forefoot (80). In saying this, it is important to recognise that not all inefficiencies will arise from external restrictions; some individuals may be inherently more inefficient by nature (35, 49, 81). This was alluded to in the sections above.

10.3 A history of 'abnormal foot mechanics' and injury risk

Abnormal foot movements were originally theorised as the cause of overuse injury in the lower limb (22, 82). This theory has high face validity: if there is an abnormal movement occurring at the foot, which is the initial link in the lower body kinetic chain, then this would place unaccustomed stress on the structures linked to them, i.e. the ankle, the shank, the knee and so on, and therefore potentially lead to injury. Thus, a significant focus of research and practise in this area of injury prevention has been the identification of mechanical abnormalities at the foot (83-85). The most frequently cited foot mechanics identified as 'abnormal' are considered below.

10.3.1 Abnormal foot pronation

Abnormal foot mechanics, specifically pronation, were first implicated in overuse injuries among runners in the early 1970's (86). 'Abnormal' was the classification given to an excess amount of foot pronation, and was clinically linked to many injuries including: shin splints, medial tibial stress reactions, lower limb stress fractures, and chondromalacia of the patella (86). Excess pronation was seen as an injurious movement pattern as it was believed that the 'normal' foot should not have to move in this way to generate propulsion. Recent literature still identifies that a high level of pronation presents an elevated risk of injury for an individual: Carr (87) found that groups with shin splints and plantar fasciitis had increased pronation at the subtalar joint. However, it was also recognised that other inefficient or

unnecessary movement patterns and foot postures present similar problems (87), as discussed below.

Advancements in movement tracking technology have allowed a much more detailed understanding of how the foot moves during locomotion. As a result, it is now recognised that pronation is not necessarily an injurious movement; indeed, those who more rapidly pronate during running activities were found to have a *reduced* risk of injury (88), arguably because this allows forces to be absorbed over a longer period of time by a greater surface area. In addition, an increased understanding of the foot and its complexity now allows us to recognise that the amount of subtalar joint pronation is determined by the range of motion available in the joints and by bony congruency (89). A foot with less flexible joints is not as capable of pronation, and therefore may not be as effective in transmitting forces to the linked structures, such as the muscles in the leg, which are the main attenuators of impact forces experienced during locomotion (90). This may be why some people with stiff, high arched feet have a greater risk of stress fractures (90), which is at odds with a theory of excessive pronation being associated with lower limb structural disorders.

10.3.2 Other abnormal foot movements

Conversely, excessive eversion of the foot during stance is also a research focus in the injury prevention literature (82, 91). Based upon this, it can be seen that both excessive eversion and pronation can potentially be linked to the development of overuse injury (22, 82). In addition, different types of foot strike patterns suggest that many varieties of gait patterns can all have potentially injurious components: rear foot strikers experience higher levels of peak forces during gait due to the nature of foot strike, mid foot runners experience greater eversions, and forefoot strikers have low levels of ankle stiffness (46). Thus, the outline of the nature of foot strike in general is relevant, as it can provide information on the aetiology of injury.

10.4 The ankle joint and injury occurrence

The ankle has been the focus of many investigations regarding lower limb overuse injuries. This research is fuelled by the substantial injury rates within various national militaries, the single most common injury in the US Military being the sprained ankle (92). The ankle is a primary joint in the kinetic chain because it links the foot to the lower limb, and therefore is highly influential to bipedal activity and physical demands. In addition, the ankle supports more weight per unit area than any other joint in the body (92), and therefore it is considered essential that it functions correctly to avoid lower limb injury. The relationship between the ankle joint, ankle sprain injuries, and other overuse injuries of the lower limb is not yet completely understood; however, investigations thus far have typically referred to two potential areas of problem with the joint: functional and/or structural instability of the ankle, which has been briefly considered before, and is discussed next, and decreased proprioceptive ability, which is discussed within the more generalised section following (93).

10.4.1 Instability of the ankle

Instabilities of a joint were defined by Dayakidis (76) and Keller (73) as a self-reported tendency of the joint to give way without exceeding the normal range of motion. Instability is currently considered to be a major cause of adverse function of the lower leg: instability of the ankle joint specifically has repeatedly been cited as a predisposing factor in the development of injury (76), and thus preventive research has primarily focused on measurement and categorisation of the degree of stability (94). The assumption underlying such work is that if instability can be successfully assessed and 'graded', level of injury risk could also be determined based on this score. This approach is based on evidence, which shows a clear association between overuse type injury and instability (94).

In order to measure the degree of instability for the purposes of risk assessment, various forms of passive testing (such as stress radiography) have been suggested. However, these

have met limited success to date (75, 95); such passive testing was found by Chen (95) to be unreliable as a measure of instability, as the force applied as part of the assessment cannot be successfully directed to a specific structure. In addition, the validity of one such test, the anterior drawer test, has been questioned as it was found that it did not represent the actual motion of the talus relative to the tibia (75).

As an alternative, instability has also been assessed based on the degree of joint stiffness, with a more rigid joint demonstrating a lower level of instability (96). However, stiffness of the ankle joint has itself also been implicated in overuse injury (97). Although stiffness may be thought to reduce the likelihood of suffering from an ankle sprain, paradoxically increased ankle joint stiffness may predispose to certain types of injury. A stiff ankle joint will transmit forces directly into the kinetic chain; while in contrast, some degree of joint laxity would arguably allow forces to be more effectively attenuated and more safely transmitted. This illustrates the complex, multi-faceted nature of injury risk, and supports a 'U shaped' risk phenomenon where extremes of any function are likely to be detrimental (12).

10.5 Decreased proprioceptive ability

Proprioception is defined by Lephart (93) as the ability to sense the position, location, orientation and movement of the body and its parts. Along with the somatosensory system, it controls for excessive strains caused by passive movements of joint structures (93). In addition, it refers to the awareness of the nature of joints and postural awareness (93). Good levels of proprioception are needed to ensure that reflex responses are sufficient, and that potentially injurious ranges of motion are avoided. This knowledge is the basis for the clinical use of proprioceptive neuromuscular techniques for promoting return to play in sports rehabilitation, and to encourage optimal function (98). If there is inadequate proprioception at and around a joint, there will be the risk of insufficient control on excessive strain; if the

musculoskeletal system cannot control for excessive strain, injurious movements can potentially occur. This can be a chronic or acute occurrence, as reflex responses are also required in order to keep the joint safe from direct injury, and a decrease in proprioception would lead to a delayed, insufficient, or absent reflex response (99, 100).

Proprioception and instability are linked, at least from a clinical perspective, with many practitioners citing or describing a deficit or reduction in proprioception as being the cause of chronic and functional instability (94). This is supported by the findings of ankle taping studies. Strapping of joints has the primary putative purpose of improving stability: using tape alone produces an activation of the local proprioceptive receptors of the skin, which is important as these cutaneous receptors and localised mechanoreceptors within muscle tissues play a key role in joint proprioception (99). Taping has been shown to reduce acute ankle injuries and improve proprioception (as tested by joint angle reproduction) (94, 99), which is proposed to be due to amelioration of the sensory deficit in both injured and dysfunctional muscles (101).

The knowledge that an injured soft tissue structure has reduced proprioception can help to explain why those who have been injured previously are at increased risk of re-injury. Following injury, the decreased sensory input from joint receptors can lead to body positions that are not ‘normal’, and a diminished reflex response (93). This may elucidate the relevance of proprioception to injury risk. In addition, it has been noted by researchers that proprioception could play a greater role than pain avoidance in the prevention of injury, and – alternatively – in the development of chronic injuries (93).

As proprioception has been frequently linked with overuse injury, there have been a number of studies which have aimed to quantify risk based on the proprioceptive ability of individuals, assessed using various tests. Jong (102) outlines two of several variables involved in proprioception: movement detection and movement discrimination, and further

states that no single test can be adequate to describe the proprioceptive ability of an individual. Thus, performance on such tests should be recognised as test specific. Movement detection is a measure, and performance improvement marker, of kinaesthesia (knowledge of joint motion (93)), and can be reliably assessed from the threshold of a slow, passive movement (103).

Building on this, a frequently used test for movement discrimination is joint angle reproduction. For example, Eils (104) used a commonly cited method of ankle angle reproduction ability. This test has been shown capable of detecting change, as proprioceptive training has been shown to increase the score of such tests (104). Medio-lateral postural sway during standing and single leg balance (known also to be impaired for those with functional instability) also shows similar improvements after proprioceptive training (73, 99, 105), and increased scores on such tests are strongly correlated with reduced injury rate (104).

Aside from joint angle reproduction, movement discrimination is also monitored through tests of reaction time, and joint position sense (JPS). Tests for JPS are typically assessed by the amount of error involved in reproducing a joint location and orientation, either actively or passively. Active JPS tests have been found to be optimal for detecting exercise related improvements in proprioception (moving through full ROM then stopping at the required angle (103)): Munn (105) found that JPS testing was able to show differences between those who did and did not sustain an ankle sprain, and those who had a functional ankle instability. This is an interesting finding as the majority of investigations of other putative factors have proven fruitless: for example, reaction time of the supporting muscle groups, such as the peroneal muscle did not show a relationship to overuse injury (105). Despite this, it is still difficult to detect minimal changes in proprioceptive ability due to a lack of valid and reliable methods of quantification (103).

The relationship between proprioception and injury (or re-injury) is so well accepted, that proprioception is generally considered the rehabilitation 'target' for trainers and physiotherapists following injury (101). Not only does proprioceptive training (for example, using a 'wobble board') help to prevent further injury by improving the participants' abilities in aspects such as JPS, it further helps with developing the strength of the tissue (through active contractions made during such training), therefore limiting the risk of potentially injurious movements (35). Given the recognised importance of proprioception in joint movement and the prevention of further injury, the use of proprioceptive test components within a given screening protocol pre-participation would seem to be justified.

As noted above, it is important to acknowledge that proprioception forms only one part of the overall risk assessment for overuse injury. Jong (102) has indicated that no proprioceptive test can in isolation identify between those who will and will not sustain injury, as there are other factors present. Beyond this, it is important to note that proprioceptive training has been reported to produce no improvements for those who do not have a deficit, and therefore proprioception training alone cannot reduce the injury risk in these individuals (106). Other potentially promising areas of investigations for injury prevention include; fitness, BMI, sex, ethnicity, and arch height; all of which are reviewed below.

10.6 Physical and anthropometric risk factors

There have been several attempts to identify an alternative, potentially anthropometry-based screening method for risk of overuse injury for use within the military. Anthropometric areas of interest have included: arch height (107, 108), height (109), weight (109) and BMI (49, 109); other physical factors have included sex, fitness, and ethnicity (109). Such investigations have yielded conflicting results, with cumulative results being inconclusive (49, 91, 110, 111).

10.6.1 Fitness level

Fitness level is one characteristic which has generated interest, and has more often been identified as a risk factor (109) than not. As all military services have a minimum fitness requirement for enlistment, it may be logical to assume that this specific risk factor is of limited relevance when considering risk factors among military personnel. However, some studies have contradicted this assumption (65), suggesting that a revision of the level of entrance fitness may be required.

10.6.2 Sex

It was a long held assumption that females were more likely to sustain an injury than their male counterparts; this has since been disproven (109), as when data for males and females were matched for anthropometric variables and fitness, there were no significant differences in the overall level of injuries. However, there are some interesting differences between the sexes which are worth noting, particularly regarding specific types of injury: for example females are more likely than males to sustain injury to the lumbar spine (109). Based upon this finding, differences in required entry level fitness between males and females, which are present in the New Zealand army, could be an influencing factor in the differences in injury rates between sexes. This theory has not been researched to date.

10.6.3 Ethnicity

Little is known about the relationship between ethnicity and overuse injury in the New Zealand army: this has not been specifically investigated or reported to date, despite health data from the wider population showing marked differences between ethnic groups (112, 113). Such a lack of focussed research is surprising, as around 40% of the army is made up of specific ethnic groups, while in the general population these groups comprise less than 20% (114). Overseas studies which have investigated the relationship between other ethnic

minorities and injury occurrence include that of Blacker (65), who found that independent of gender and fitness, white people in the US army were more likely to sustain injury than any other of the ethnic minorities. This would appear to justify a similar study being done in NZ.

10.6.4 BMI

Increases in BMI within non-military populations have typically been shown to cause increased risk of injury, especially during weight bearing activities (109). In contrast, studies within military settings have found that an increased BMI is associated with *reduced* injury risk (65). The most plausible explanation for this is that in situations involving load carriage, relative load carried is decreased with increased BMI, and therefore the absolute weight has a proportionately less negative affect for larger individuals (65). It appears then that BMI has a complex relationship with injury risk, and this is something that should be addressed in future investigations.

10.6.5 Arch height

Despite long-standing 'flat foot' assumptions concerning arch height and injury risk, research findings in this area have been contradictory. In some research studies, there was no difference found between those who were, and those were not, susceptible to overuse injury based on static angular measurements of the foot arch (115, 116). Further to this, arch height was found not to correlate with pain or dysfunction (116). However, recent research by Riskowski (117) which used a considerable cohort (N=1856), and follow-up period (3 years) identified that a planus foot structure was associated with increased odds of knee and ankle pain, whereas cavus feet were at increased odds of ankle pain. The occasional associative results seen in the literature between flat footedness and injury can be explained – at least in part – by gait mechanics. A flat footed person typically has an excessive range of motion (ROM) during gait (82, 118). Excessive ROMs at the joints place the individual at a greater

risk of overuse injury (119), as will be discussed in depth below. This is not always the case, but does explain why there are limited reports of correlations between flat feet and overuse injury.

The injurious association with *dynamic* measurements of arch heights indicate that notably higher navicular drops (where the magnitude of navicular drop is not correlated with habitual arch height) are seen in those who develop medial tibial stress syndrome (MTSS) (120). This is logical as the movement of the arch causing a navicular drop inevitably would place stress on the muscles which maintain the arch, which originate from the tibia. This evidence regarding dynamic changes in arch height could potentially be incorporated into a screening protocol.

10.6.6 Summary of risk factors

It seems there are many risk factors, from a variety of sources, to consider when investigating the aetiology of overuse injuries. As the military is a special interest group for the current investigation, the following section considers the unique injury risk environment provided by military service, and discusses the issues which arise from overuse injury research within such an environment.

10.7 Injury occurrence in the military

Injury occurrence in the military is known to be associated with both intrinsic and extrinsic risk factors. Leggat (121) proposes that the intrinsic risk factors for lower limb injuries in the military include: previous injury, psychological make-up, hyper-pronation, and tibial bone width. Extrinsic factors include: ground surface variations, training in a fatigued state (35), and type of footwear (65). Both extrinsic and intrinsic risk factors will be discussed here.

10.7.1 Unique character of the military training environment

From the outset, it would seem that service in the military is potentially associated with the majority of commonly identified risk factors for overuse injury of the lower limb. Despite many researchers considering the athletic, running, and military populations as interchangeable, and thus findings generalizable between the groups with respect to injury prevention (122, 123), there are a considerable number of notable differences. As will be discussed in depth later in this chapter, there are a limited number of variables which have been consistently proven to be a risk factor for overuse injury of the lower limb, these include load carried and the restriction of arm movements during bipedal motion (64, 70). Both of these variables relate a great deal more directly and specifically to military groups than runners or athletes: military service (especially in the land forces) includes regular carrying of loads, long distance (or 'forced') marching, uneven terrain, and the restriction of arm movement through weapons carriage. Military service therefore imposes a unique environment contributing to the prevalence of injuries.

This injury risk is held in common by defence forces throughout the world. In the British army, musculoskeletal injuries, including overuse injuries, are highly prevalent (up to 11% during initial training), posing a high cost to the soldier, the army, and the state (65). Here in New Zealand, Davidson (35) and Baxter (1) both found that 10% of the NZDF at any one time were affected by lower limb overuse injuries. Of these, 35% were ankle sprain, and 16% were knee sprain.

In the USA, approximately 50% of Marine Corps training recruits sustained an overuse injury to the lower limb (109). Of these, the lower leg was the most common site, followed by the knee, then the hip joints. Further to this, it is known that the single greatest cause for lost training days in the US Marines is stress fractures (108).

The injury rates described are not unique to the Marine Corps. No difference has been identified between the US forces, whether air force, navy or army personnel: all have similar high levels of overuse injury rates (111); approximately half of all injuries sustained by personnel from the US army were overuse type injuries, the majority (82%) of which were located in the lower limb (124). Typically, these are made up of: 22% located in the knee, 13% in the ankle or foot and 20% in the hip, pelvis, back or spine (based on an evaluation of 1.6 million injuries) (66).

The explanations for such a high frequency of injuries are likely to be multifactorial, as the training environment is necessarily harsh and physically demanding. The various sources of increased risk likely to be associated with this training environment, namely the equipment and aspects of physical training, are considered in more detail below.

10.8 Military footwear and physical activity

Both footwear and unaccustomed physical activity have been previously identified as the two greatest precursors for foot fractures in the military (110, 125). Unfortunately, such risk factors are unavoidable during military service, as they are involved in daily training and active duty. The injury risk associated with change in footwear (principally during initial training) is exacerbated by the necessity to wear a military style boot, which provides substantially reduced shock attenuation compared to the average casual shoe (91). This then means that shock is directly transmitted through the lower limb and must be absorbed by the joint structures. Adding to this, the military boot restricts movement of the foot to the extent that the loading placed on certain structures is excessive and therefore can lead to tissue break down (91). These two topics (footwear and physical activity) will form the main topics for discussion in this section.

10.8.1 Boot ergonomics

Many would argue that abnormal mechanics of the lower limb are the result of wearing sub-optimal footwear, and that a return to barefoot walking could potentially solve many of the current injury related problems. This has generated clinical and research interest encouraged by the biomechanical analysis of barefoot walking. The differences in mechanics associated with barefoot walking include a significant displacement of forces to the rear of the foot; a significant reduction in time spent in all phases of ground contact; and a significant increase in the time spent during toe-off when compared to shod gait (126, 127). Whether or not these changes are optimal for function, and thus could potentially reduce injury rates, remains to be seen. The practical implications are however clear: potentially injurious surfaces make widespread adoption of barefoot walking impossible, for both civilians and the military.

The selection of adequate footwear has always been a challenge for the military, in New Zealand as elsewhere. A focus on footwear as a primary intervention might appear to be justified, given the knowledge of how footwear can affect injury risk: for example, Grier (97) found that footwear choice had implications for shock absorption, arch support, navicular fatigue, plantar pressure distribution, energy cost of movement, and fatigue. An ideal choice of footwear procurement may be regarded as a "*Catch 22"*; the high level of training and function undertaken requires a highly specialised boot, ideally, one that is individually matched to the soldier. For reason of logistics and cost (New Zealand army representative, personal communication), it is however impracticable to provide each soldier with a customised boot.

Different foot types and gait styles do however mean that one style of boot will never be optimal for every member of the military. Internationally, there have been various attempts to identify the optimal style of boot for the services, but to date this appears to have been met with little success with respect to overuse injury prevention (1, 35). It seems that no

matter which boot is chosen, there will be a substantial portion of personnel who will describe being adversely affected by the choice of boot. As a result, in some circumstances, personnel are permitted to procure their own footwear, which best suits their individual needs. An innovative approach to supplying one boot for all, and still allowing for customisation, was recently completed by the Australian army. The results of the trials as summed up as follows.

An all-weather combat boot was designed and called the 'Trekka™'. This was chosen by the Australian army due to its innovative design and cost-effectiveness for the army's requirements. The boot was different from previous issue boots, as it was essentially built around a customised interior. With a choice of inserts the boot could, in essence, be made to suit the foot in question. In theory, this would have addressed the problems associated with standardised 'one type for all' military issued boots. When the boot was trialled however, there were significant problems. The choice of inserts available turned boot issue and replacement into a major logistical challenge. Stores of all the various options for interior components had to be made available, which limited the cost-effectiveness. In addition, soldiers were not educated on what a well-fitting boot would feel like, and neither were the staff who supplied them. The result was that very few soldiers ended up with a well-fitting boot, and the problem with injuries did not change (Australian army representative, personal communication).

Similarly, Grier (97) investigated the use of a more mechanically advanced boot sole, which had increased shock absorption, cushioning, and ventilation. This also produced little difference in terms of injury prevention, the benefits being limited to the reduction of injuries directly to the foot. The reasons for these apparently disappointing results are unclear, but speculatively, they would be similar to the issues identified for the Trekka™ (i.e. lack of understanding and resourcing challenges). This problem is likely not unique to the Australian

army: the logistical complications and associated costs are a universal issue with such a high-tech boot. Challenges of soldier education and supplier knowledge would be comparable in other armies, including the New Zealand army. Thus, this approach does not seem to provide a practicable solution to the problem of lower limb overuse injuries.

10.8.2 Standard issue footwear

Baxter et al (1) recently identified that up to one third of New Zealand soldiers were wearing footwear that was the wrong size. This was in keeping with previous anecdotal reports from colleagues within the NZDF (Personal Communication). Similar problems have been reported in the US army: Teyhen (107) identified that 35.1% of basic trainees were wearing a boot which was a full size too big or too small. This highlights two important points: firstly, those persons responsible for issuing footwear are not specifically trained to do so, and secondly that soldiers entering basic training do not identify that they have been supplied with ill-fitting footwear. An additional problem in the New Zealand army is that the boots were not made in sufficient ranges of shoe widths to accommodate variations in foot anthropometry amongst the soldiers. In some cases the boots worn were up to three sizes too large (long) in order to allow sufficient width (80).

Similar problems have been reported for running shoes issued, where matching of foot type to running shoe types was less than 44% accurate in the US army (107). This is an interesting finding, given the challenges experienced by both the US and NZ armies, resulting in the need to allow for the purchase of non-standard (non-issue) footwear. Anecdotal evidence has suggested that doctors within the New Zealand military write on average three prescriptions per day for non- issue running shoe footwear (Personal Communication). This evidence suggested that soldiers declined the available issued running shoes because they were not ‘trendy’, and simply used the tried and proven excuse of having sore feet or knees in order to be given a prescription for a more acceptable alternative (Personal Communication).

In contrast to this, the current evidence may indicate that such soldiers did indeed require alternative running shoes, as they were unable to find an issue shoe which was appropriate for their foot shape needs or movement style.

In keeping with uniformed services internationally, the process used to issue footwear by the NZ army would appear to be generic and overly simplistic, given the anthropometric variation in the population of recruits and service personnel. It is also performed by staff who do not have sufficient training to undertake what is a complex process. It was noted during previous research (1) that army recruits are simply asked to identify the size of boot they needed. This is problematic for several reasons: firstly, many recruits genuinely don't know what size foot they have, have never been properly assessed and fitted for footwear, and even if they look at the markings on their existing footwear, it is not always clear whether the size indicated is a US or a UK size. Related to this are differences between manufacturers or brands: manufacturers have their own specifications of sizing cut-offs, and thus a size US9 in one branding isn't necessarily the same size as a US9 in another branding. Given these issues, there is a clear need for a standardised, simple, but effective measurement system for those who issue footwear to the military.

While this may appear straightforward, implementation of such a process is challenging. While it is widely recognised and acknowledged that the US military dedicates a substantial amount of resources (i.e. proportionally more time and money than other national armed/uniformed services) to the development and implementation of footwear fitting procedures in the military, the evidence suggests (conservatively) that systems used in the US Army also got it wrong in more than one third of cases (107). There are, therefore, good reasons to suggest that procedures for footwear fitting in the military should be standardised and 'professionalised', also operationalized in such a fashion that stores personnel can understand. A calibrated foot device should be available, and a monogram consulted for

accurate fitting, as is currently carried out in commercial footwear practice. There is also a need to increase the skills and training of relevant military personnel on the importance of accurate fitting of footwear for individuals, including its relevance to care, maintenance, and injury reduction techniques.

Given the evidence regarding injury prevalence, incidence rates, and the time and effort spent trying to prevent or predict lower limb overuse injuries discussed here, it is somewhat surprising that there has been relatively little focus on the provision of correctly fitting footwear for military personnel.

10.8.3 Physical training

Physical weight-bearing exercise which is integral to military service has been implicated in overuse injury development (128), this is exacerbated by the carrying of weight (e.g. in a backpack) and a rifle (64). Increased training 'load' (where 'load' is a combination of frequency and intensity) is also an unavoidable necessity in military service. The sharp increase in physical activity levels experienced as part of basic training (where new recruits are initially trained) has often been cited by practitioners (Lt Col Dunn, NZDF, personal communication) as the single most influential factor in the resultant overuse injuries (125). However, the nature of military training clearly necessitates a degree of physical fitness, and as such, the manipulation of physical training load continues to be a contested topic within research.

Basic training recruits experience the highest incidence of overuse injuries of any group within a military force, arguably because of the cumulative effect of changing footwear to military boots for the first time, and the sharp increase in training level which coincides with it (22, 111). The challenges associated with the extreme physical training demands accompanying military service will be discussed in more depth later below (refer to the discussion titled "*Overuse injury prevention*"). This notwithstanding, the wearing of boots

and completion of high levels of demanding physical activity would appear to be necessary and unavoidable when in the military.

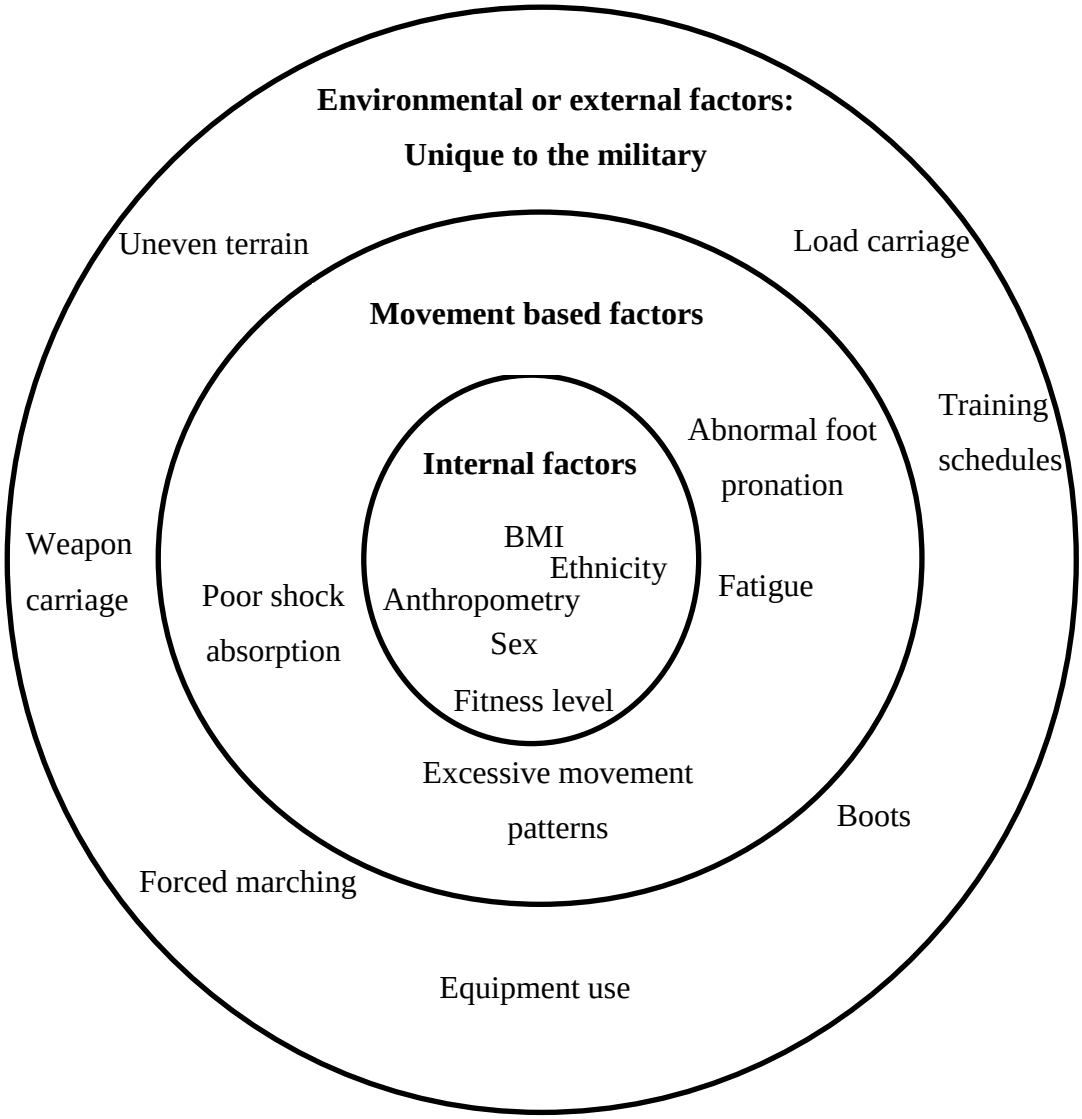


Figure 1: Diagram to illustrate the potential sources of overuse injury risk, as discussed in the current Chapter.

10.9 Final comments: injury aetiology

Acknowledgement of the risks of overuse injury associated with military service as highly problematic does little for their prevention. The magnitude and impact of the problem, highlighted in the incidence rates, is exacerbated by the costs involved in the treatment of

such injuries. Among New Zealand (NZ) army personnel, there is a 10% prevalence rate of these potentially preventable injuries (1, 35). If a soldier is injured, and they cannot participate in daily training, or serve on a deployment, associated costs are in addition to those related to treatment or clinical management. With approximately 4500 personnel serving in the New Zealand army, approximately 450 personnel are injured at any time. Assuming that these individuals could not work on the days they were injured, and we account for paid sick-day costs only, this represents a cost of 450(people) x 365(days per year) x $90(approximate gross daily wage) = $14,782,500.00 annually (1).

Based on this, it would seem that even though the inherent risks of military service are high, preventing such injuries has the potential to save the military (and therefore wider society) a substantial amount of money, time and other resources. Injury prevention within the military including the findings of current literature is discussed below.

10.10 Overuse injury prevention through risk factor identification

Research literature on the prevention of lower limb overuse injuries is considerable (129-131). Typically, the focus of these has been investigation of potential variables as risk factors of injury (49, 65, 110). Much of this literature utilises military groups as the participants, presumably due to their predictable lifestyle and the relative ease of data collection and follow up (65, 110, 130, 131).

Both internal and external factors contribute to overuse injury risk and occurrence (49). External factors range from equipment choice (as discussed above) to daily weather differences: Beynnone (49) found that the majority of injuries among college athletes occurred when the running surface was dry and the weather was hot. The nature of most social, economic, and vocational environments mean that few external factors can be controlled or minimised: life, activities, work, and sport for the main part cannot stop because of weather, and most of the equipment which can potentially contribute to injury risk, such as

footwear or carrying weight, cannot be discarded as these represent an integral part of today's world. However, such equipment can be made as effective as possible for the environment.

Internal factors for overuse injury have been investigated extensively, as is discussed below. Certain anthropometric and functional variables have had considerably more attention in the research literature than others, including the ankle joint as a focus point, and the individual's proprioceptive ability. The findings of current literature relating to risk factors are presented below; the more saturated areas are discussed separately.

10.11 Military training

Two factors which have been investigated to date with consistent findings regarding injury risk are absolute load carried, and the carrying of a rifle (64). Absolute load carried increases ground reaction forces, medio-lateral forces, and time spent in stance, which are strongly related to changes in gait (64). An increase in ground reaction forces, both loading rates and impact peak, also show a significant correlation with increased injury rates (88). In addition, when load is carried on the back (in a pack) forward lean is increased, causing an increased ground reaction force (GRF) at touchdown and longer stride lengths (64). Touchdown is arguably the most significant phase of the gait cycle with respect to overuse injury, as the positioning of the lower limb causes breaking forces (i.e. 'slowing down' or eccentric movements) which result in shear forces transmitted through the limb; the use of back packs result in the highest braking forces (70). The reason that carrying a rifle exacerbates injury occurrence would appear to be associated with the restrictions placed on arm swing during gait, the kinetic result of which increases medio-lateral movements of the foot and therefore a reduction in stability (64).

Unfortunately, both of these are essential components of serving in the military and therefore cannot be changed within this group. Anthropometric characteristics discussed previously are also essentially non-modifiable; e.g. we cannot change the arch height of an

individual. Thus, military-centred injury prevention research has recently directed interest towards the use of interventions as an alternative means of reducing overuse injury rates.

10.11.1Intervention strategies for injury prevention

For any injury prevention intervention measure to be effective there must be acceptance, adoption, and compliance from those who have been assigned the intervention (132). Leggat (121) conducted a comprehensive systematic review which aimed to identify through meta-analysis the best intervention for reducing musculoskeletal injuries among military personnel. As risk factors for overuse injury are not always known, or modifiable, some researchers have investigated the potential of employing an intervention instead. Based on the findings of Leggat (121), the primary recommendation to reduce injury was to incorporate individualised and specialised training programmes, however, within a military setting, this is not feasible. Instead, perhaps the most feasible and plausible recommendation made by Leggat (121) was that specific biomechanical conditions of the individual's performance should be identified, and based on these, movement corrections made as required.

Various injury prevention interventions have been trialled in the military, including: stretching programs, concurrent preventative training, and modifications to physical training. Findings from the most salient studies are outlined below.

10.11.2Stretching programs

Herbert (133) conducted a systematic review which investigated the potential effectiveness of various types of stretching programs (both before and after exercise) in reducing injury occurrence. There were no differences seen between groups who stretched and those who did not. This result was consistent for all forms of stretching, including dynamic, static, and proprioceptive neuromuscular facilitation (PNF), and across various dosages and regimes.

10.11.3Physical training

Leggat (121) conducted a systematic review which investigated reduction of physical training and its effect on overuse injury occurrence among military personnel. It was found that reduction of training by 60% produced a 40% reduction in overuse injuries. Subsequent studies have also identified similar correlations between injury rates and physical activity (132). In New Zealand, most recruits come to their military basic training relatively unconditioned physically, which makes them subject to high levels of fatigue during their intensive training course. These circumstances clearly have potential to cause overuse injury in these recruits.

Given the nature of the military training environment, and the requirements of being a soldier, reductions in training as explored by Leggat (121) are not feasible for the group of interest. As indicated, carrying weight and moving quickly on foot are also essential skills needed for all soldiers in any army. Despite this, the NZ army has in recent years reduced its physical training load over basic training in an effort to reduce the high injury rate. Results observed here were similar to those seen when the US army employed the same strategy (124): a marked reduction in total injuries being referred to the physiotherapist or other medical services. This notwithstanding, injury rates remain a significant problem in the military, including the NZDF. Thus, despite such encouraging results, it would not be a realistic option to reduce training loads further, and therefore this prevention method lacks feasibility (22).

10.11.4Concurrent training

Brushoj (132) investigated the potential of a concurrent preventive training programme for overuse injury reduction. The programme included aspects of strength, flexibility and proprioception, yet it did not result in a reduction in overuse injuries among soldiers (N=1020). Conversely, Davidson (35) reviewed the research for such interventions and found

that stability-based training programmes were effective in reducing specific injuries. Unfortunately, the practicality of such a course of preventative treatment remains limited, particularly as there is little spare time available for this in basic training courses. In addition, the effectiveness of such programmes have been found to be critically dependent upon individual compliance (132), which cannot always be guaranteed among the groups of interest.

10.12 Summary on interventions

To summarise the discussion on interventions, the main problem with current literature concerning risk factors and interventions is the lack of conclusive or consistent results with respect to injury reduction. Because an effect on overuse injury rates is occasionally observed, it may indicate some level of association between the variable of interest (e.g. concurrent training or stretching programs) and injury prevention and management, but findings are not consistent. Despite this, it is now becoming increasingly accepted by practitioners and researchers alike that risk of musculoskeletal overuse injury is much more complex than simply anthropometric or physical variables, and is related to the putative presence of several key characteristics and factors that may not be consistent for all people at risk (67). It is now recognised that one factor alone cannot comprehensively account for injury risk (which has historically been the most frequent assumption underlying scientific investigations). The current thesis acknowledges this, as the injury risk assessment protocol being researched is comprised of three functional assessments.

10.13 Limitations within the literature

The reported limitations of previous studies include their apparently poor understanding of the contemporary knowledge base, selective reporting and overly limited methodologies (134). For example, despite the overwhelming majority of literature which has found no

differences between arch heights for injury risk, studies continue to be published in recent years which investigate this as a possible risk factor (107). To continue to replicate previous studies in anticipation of achieving different results would appear unjustified, and in some circumstances this could potentially be unethical.

The thrust and direction of previous literature may also be considered as fundamentally flawed: the majority of studies have focussed on idealised 'one approach fits all' interventions. The justification of providing en-masse interventions is not clear, and the cost-effectiveness is, at best, questionable. The use of such en-masse intervention methods is likely to be linked to another significant challenge of injury prevention: that our knowledge of risk factors is still incomplete. Therefore, an improved understanding of such risk factors is needed to better inform the direction of prevention studies. This has informed the main direction of the current study, which is concerned with justifying and validating risk factor assessments, through an investigation of the psychometric properties of three visual assessments of injury risk.

10.14 In the following Chapter

The current Chapter provides an overview of overuse injury development and risk factor assessment, focussing on the military population. The next Chapter is a narrative review of current research which investigates the psychometric properties of visually-based functional screening tests.

11 Chapter 3. A review of functional screening tests

11.1 Overview of Chapter

The previous Chapter outlined overuse injury aetiology, injury prevention and potential interventions. The aim of the current Chapter was to review the current evidence of psychometric properties of visually-based, functional screening assessments which identify risk of overuse injury. The following begins with a brief history of clinical screening assessments. This informed work reported in subsequent Chapters (5, 6 and 7) as it identified the gaps in current knowledge, and allowed informed progression of a functional screening protocol based on the strengths and limitations of those previously tested. The research question addressed here was:

What is the current evidence for psychometric properties of clinical assessment tools for overuse injury risk identification?

The structure of this Chapter takes the form of a narrative review of currently available literature.

11.2 Background

As indicated in Chapter 2, injury prevention strategies are essential to address the high prevalence of lower limb overuse injuries within society, and the associated costs. Primary prevention strategies involve identification of potential risk factors (as discussed in Chapter 2) for overuse injury through the development of appropriate screening tests. Many screening tests are available which screen for chronic injuries of the lower limbs (68, 135-137), low back (138) and associated neuromuscular deficits (136), and these are documented in clinical reference books readily available to the practitioner (14).

The purpose of this review is to investigate what is currently known regarding the psychometric properties of current injury screening protocols which identify overuse injury risk in the lower limb or a neuromuscular deficit which is associated with such injuries. This review included studies of any psychometric property, i.e. reliability (both within and between observers) and validity (criterion based, test, or construct) for each screening tool.

It is important to outline what is meant by the reliability and validity of a screening test, and why these aspects are important to the interpretation of the results we get from such a test. The reliability of a screening test (or any test) is how accurately the test produces consistent results with repeated application (139). The specific question posed in an assessment of reliability is: if we were to repeat the same test under the same conditions, with the same observer and/or a different observer, would we get the same result? This is typically investigated in two parts, both of which are equally important in the specific example of clinical screening tests: inter-rater reliability and intra-rater reliability (140).

Inter-rater reliability (also referred to as between observer, inter-tester, or inter-assessor reliability) refers to the extent to which different individuals can produce the same test result under the same conditions (140). This is clearly an important property of a test, especially in the field of clinical screening; a screening test result for a patient should not depend upon their choice of practitioner. For example, consider a hypothetical case of a High School rugby team that presents at a physiotherapy clinic in Dunedin, New Zealand, to be screened for ankle instabilities. On the basis of the screening protocol, four players are identified as being at high risk of injuries and therefore requiring preventive training. If the same team went to an alternate practice, and were administered the same screening test, ideally the outcome should be the same four players are identified as being high risk. If the high risk players are not the same as those identified in the first clinic, it indicates that the screening test lacks an

element of reliability and therefore creates several issues for the players and the practitioner: who is genuinely at risk, and who should be receiving preventive treatments?

Intra-rater reliability (also called within observer, intra-tester, or intra-assessor reliability) refers to the extent to which the same individual can produce the same test result under the same conditions (140). This is an equally important quality of a test, as tests are frequently repeated within the clinical setting to assess improvements or monitor changes. A high level of intra-rater reliability reduces the variability in the measurements or outcomes obtained, and therefore provides more robust data. Thus, it allows conclusions to be drawn that are not subject to error. For example, if a change is observed from pre-season to in-season where intra-rater reliability is high, we can conclude a real change has occurred. If such reliability is low, then we cannot say whether there has been a real change or just a difference occurring by error.

To summarise, under constant conditions, inter-rater reliability refers to whether clinicians or practitioners can obtain the same result from a given test, whereas intra-rater reliability is whether the same clinician can achieve the same result over multiple assessments (140).

Validity can be defined in terms of three different aspects: test (or 'content') validity, criterion based validity, and construct validity (141). Test validity addresses the question; '*Does the test measure what it was designed to measure?*' (141). Criterion based validity is based upon comparison of the results gained from the current (screening) test with those gained from a well-respected outside measure, typically referred to as a 'Gold Standard' measure (a benchmark test that is the best available under reasonable conditions) (141). Construct validity is a psychometric concept, testing whether the test is a measure of the construct it claims to measure. For example, if the measurement tool is claimed to be an indirect measure of ankle injury susceptibility, then those who were identified as being high

risk by the test should experience more injuries than those who were identified as being low risk. If this were in fact the case, then the test would be said to have construct validity (141).

The importance of assessing the validity of a test is thus self-evident: if a test lacks validity, the justification for its continued use is questionable. Establishing criterion based validity is important as it gives an indication of the comparability of the assessment against an established assessment tool. In general, a cheaper or more practical test is compared to possibly more expensive or impractical gold standard tests. In the case of hindfoot angle assessment, a goniometer assessment has been compared to radiography assessments in a number of studies (142, 143). The level of agreement between the two tests represents the criterion based validity, and indicates how acceptable the use of the alternative is compared to the gold standard (as defined by (141)).

11.3 Search strategy

While this was not designed as a systematic review, a designated search strategy was employed to identify relevant papers for inclusion in the review. The following databases were searched, from inception to the date indicated.

Medline (Ovid) (18/12/13), AMED (18/12/13), Cochrane library (18/12/13), PEDro (18/12/13), CINAHL (18/12/13). Hand searches and reference list searches were also performed, based upon relevant journals and reference lists of retrieved papers.

The following keywords were used:

Reliab*, valid*, musculoskeletal, clinic*, screen*, injur*, chronic, low* limb, low* back, neuromuscular.

11.3.1 Inclusion criteria:

1. The article is concerned with a functional (movement based), clinically-based injury risk, or dysfunction, assessment for the lower limb, or neuromuscular deficit.

2. The screening test was designed to be utilized on asymptomatic persons, for risk of future pain or injury.
3. The test of focus within the article does not require any specialist equipment or computerized software which would not be routinely available. For example, goniometers, measuring tapes and stopwatches are eligible, force plates, foot scanning devices and exercise machinery are not.
4. The article primarily investigates the reliability (inter or intra-rater) and/or criterion based, test, or construct validity of such an assessment.
5. The article reports results in the form of descriptive or correlation statistics: for example, Kappa, Spearman's correlation co-efficient, or percentage agreements.

11.3.2 Exclusion criteria:

1. The article does not provide a complete and comprehensive results set.

A total of 599 studies were identified, of which 97 were duplicates. Following the screening of abstracts, and the removal of studies which did not meet the inclusion criteria, eleven studies remained (Figure 2). Studies which investigated validity of screening protocols are discussed first, followed by reliability investigations. The final study included in the current review was the only multi-dimensional test investigated for psychometric properties found in the literature.

11.4 Results

11.5 Studies identified

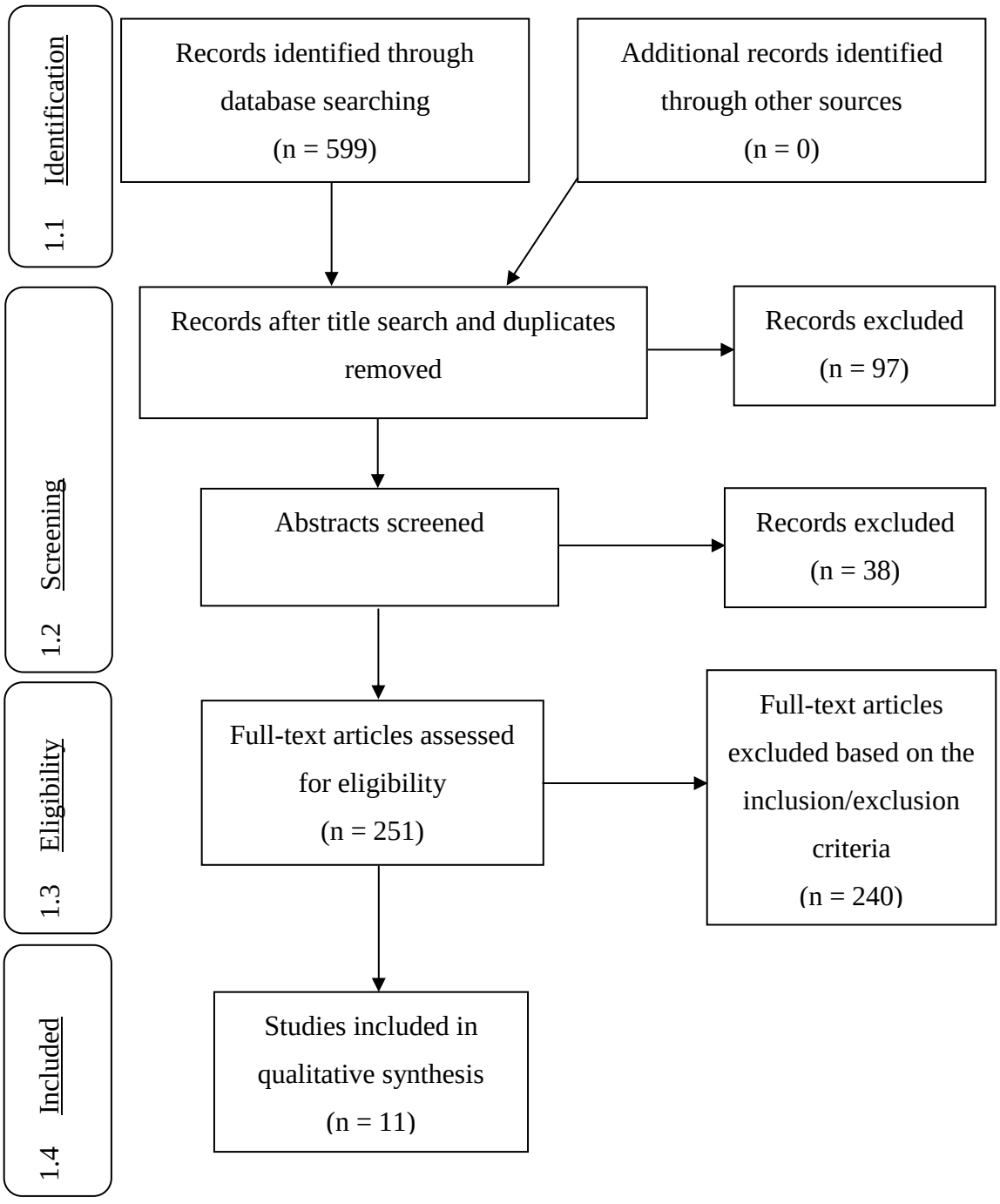


Figure 2: Diagram showing number of retained studies during the search process (number of retrieved and retained studies shown in brackets).

Table 11-1: Collected measures of psychometric properties from clinical tests of included studies. ICC; inter class correlation coefficient, K; weighted kappa score.

Clinical test	Intra-rater score	Inter-rater score	Criterion based validity	Construct validity	Test validity
Multi hop test (144)	ICC >0.9*	ICC >0.9*	n/a	'supported*'	n/a
Gillet test (137)	K <0.08	K <0.08	n/a	n/a	n/a
BESS (145)	ICC >0.50	ICC >0.46	n/a	n/a	n/a
Hip extension test (138)	n/a	K = 0.72 to 0.76*	n/a	n/a	n/a
Hindfoot classification (two studies) (68, 142)	ICC = 0.88 to 0.90*	ICC = 0.56 to 0.65	48% ICC = 0.64 to 0.95*	n/a	n/a
Single leg balance test (146)	n/a	K = 0.90*	n/a	n/a	P <0.05*
Sport specific preseason test (67)	ICC = 0.63 to 0.99*	ICC = 0.88 to 0.97*	n/a	n/a	n/a
Star excursion balance test (135)	n/a	n/a	n/a	n/a	'supported*'
4-hop test (136)	ICC = 0.76 to 0.92*	n/a	n/a	n/a	n/a
T-test (135)	ICC = 0.82 to 0.96*	n/a	n/a	n/a	n/a
Horizontal and vertical HJAT (147)	H ICC = 0.93* V ICC = 0.92*	n/a	H R^2 = 0.62* V R^2 = 0.63*	n/a	n/a

** indicates an agreement score which met the clinically acceptable minimum level as chosen by the respective authors.*

11.5.1 Star Excursion Balance Test (SEBT)

The SEBT is an objective measure to differentiate deficits and improvements in dynamic postural control, related to low extremity injury, and thus is predictive of lower extremity overuse injury (135). Lines are marked on the floor, in the shape of a three pointed star. The participant stands on one leg in the centre of the star, with hands on hips. The participant attempts to slide their non-supporting leg to the tip of each point of the star, and is scored based on how far along the points they can reach without losing balance.

Given the popularity of the research investigations, Gribble et al (135) conducted a systematic review of current literature on psychometric properties of the SEBT.
The included studies had to be original research, written in English, where the SEBT was the primary measure included. The authors identified over a decade of SEBT research, and thirty studies which met the inclusion criteria. Based on the collated evidence, test validity of the

SEBT is supported with respect to prediction of future pain or injury, specifically: ankle instability, anterior cruciate ligament injury and patella-femoral pain.

11.5.2 The multi-hop test

The Multi-hop test is an assessment of postural control aimed at identifying those who have dynamic stability deficits (148). In the opinion of those authors (149), a dynamic stability deficit is accepted as one of a variety of risk factors that can lead to development of both overuse injuries and sprains (149).

Specifically, the test is an assessment of the ability to keep the centre of gravity over the base of support. The authors explained and justified the protocol as follows: In order to preserve stability, postural strategies are employed. These are either predictive (anticipatory) or reactive and involve either fixed support strategies (flexing the centre of mass around the base of support), or changes in support strategies (hopping from side to side or placing the opposite foot down). Fixed support is what naturally should happen during movement, i.e. making adjustments without changing the base of support. (149)

The multi-hop test consists of a single leg hop between various points on a floor. It includes aspects of lateral, medial, and frontal plane movements. The test is scored based on the number of postural corrections, for example: falling, shuffling, loosing hand placement from the hips, with a higher score indicating poorer balance. This test scores well on both intra and inter-rater reliability- ICC >0.90 for six observers who assessed twenty nine symptomatic (chronic ankle instability) and twenty nine healthy participants on two occasions (149). Construct validity was supported, as scores in the multi-hop test were reduced for those participants recruited because they experienced chronic instability of the ankle joint. The high level of standardization and the low level of complexity are specifically noted as strengths of this test which produce the high scores for inter and intra-rater reliability assessments (149).

11.5.3 Hip Joint Angle Testing (HJAT)

Muscular extensibility of the hamstrings has long been used to identify changes in lumbo-pelvic rhythm, diagnose low back pain disorders, and predict risk of hamstring muscle and patella-femoral injury (147). Both the horizontal and vertical HJAT are widely used methods in clinical and sports physiotherapy (147). Both of these tests are performed visually by the physiotherapist. Despite the assumption that either of these tests can be used interchangeably with a straight leg raise (SLR), assessed using an inclinometer, there had never been an investigation into the validity of this claim.

Fifty eight participants were assessed by one examiner on three occasions, using the horizontal and vertical HJAT, and the SLR. The intra-rater reliability for the horizontal and vertical HJAT were ICC0.93 and 0.92 respectively. Despite this positive finding, the criterion based validity comparison with the SLR produced R^2 values of 0.62 and 0.63 respectively, corresponding to only 'moderate' correlation in the scores.

The authors noted that although the HJAT did not produce acceptable validity scores, further research is essential, as clinicians need tests which produce as high as possible reliability and validity scores, while requiring minimal equipment and preparation time.

11.5.4 The 4-hop test and the T-test

Agility-based functional testing is a common means of screening individuals (usually athletes) for knee joint function, leg symmetry, prediction of leg power and strength, injury risk and for assessing return to play. Munro et al (136) identified the 4-hop test and the T-test as two of the most frequently used agility-based functional tests used by both clinical and sports physiotherapists. In their study, twenty two athletes completed the 4-hop test six times, and the T-test four times during each session. There were three sessions in total, assessed by one examiner. The 4-hop test is scored based on maximum distance achieved by four consecutive unilateral hops. To perform a T-test, lines on the floor are placed in the shape of

a ‘T’. The participant is required to move forwards, then to the left, then right, then backwards, over the lines as quickly as possible while facing to the front, and the test is scored based on time.

The intra-rater reliability of the 4-hop test and T-test were ICC0.76-0.92 and ICC0.82-0.96 respectively, and there were no gender differences in performance or reliability measures. The authors noted that the development of reliable, standardised protocols for such functional tests is important, which highlights the need for future studies in this area.

11.5.5 The Balance Error Scoring System (BESS)

Finnoff et al (145) investigated the inter and intra-rater reliability of the BESS functional test, a common test in clinical practice, and frequently used to assess concussed patients (145). The BESS is designed as a practical and inexpensive tool for the assessment of functional stability. Subjects are required to close their eyes, place hands on hips, and stand in three different positions: feet together, tandem stance, and single legged stance, on two different surfaces: firm and foam, and for twenty seconds each. Whilst the participant is performing these stances, a clinician scores their postural stability based on an objective list of errors, which include:

1. Lifting the hands off the iliac crest
2. Opening the eyes
3. Stepping, stumbling, or falling
4. Lifting the forefoot or heel
5. Remaining out of the test position for more than five seconds
6. Moving the hip into more than 30 degrees of flexion or abduction (145)

The maximum number of faults possible is ten for each part of the test; thus, the maximum possible BESS score is sixty. In that study, three assessors scored thirty participants on two occasions. Intra and inter-rater reliability was reported as intra-class correlation coefficient

(ICC) values for each of the six parts of the BESS assessment. ICC values for intra-rater reliability ranged from ‘average’ 0.50 (single leg stance on foam), to what the authors described as ‘extremely high’ 0.88 (single leg stance on a firm surface). Inter-rater reliability scores ranged from ‘below average’ 0.44 (firm tandem stance) to ‘high’ 0.83 (firm single leg stance). (145)

Finnoff et al (145) investigated a visually based, objectively measured, functional assessment of single limb stance, which makes their investigation highly similar and comparable to that which will be performed as part of the current thesis. The results which were most relevant to the current thesis were the outcomes of the assessments of inter and intra-rater reliability for the test of single limb stance on a firm surface. In addition, Finnoff et al (145) adequately powered their study (80%), which is also the case in the current thesis, and therefore comparison between the two sets of results is important.

11.5.6 Gillet test

The Gillet test is an assessment of ilio-sacral and sacroiliac dysfunction, where a positive test indicates the need for intervention. The Gillet test is a comparatively recent addition to the clinical testing repertoire (around two decades old), and has had a relatively high profile in the research literature. Previous studies concerned with the performance of the Gillet test had been somewhat inconclusive due to methodological choices, thus Meijne et al (137) aimed to provide a robust assessment of the inter and intra-rater reliability of the test for both symptomatic and asymptomatic individuals.

The Gillet test is scored on the basis of eight manual contacts within the pelvis, as the participant stands and raises a leg. A ‘positive’ outcome for each of these parameters is gained when there is an ‘adverse’ movement of the contact point during the leg raise. It was found that Kappa scores for both intra and inter-rater reliability were never higher than K=0.081 (thirty seven participants, two examiners, two occasions), and were, more often than

not, negative. The authors noted that this test, along with other similar clinical assessments, are ultimately based on examiners' individual interpretations and are thus subjective. They proposed that greater standardization, including enhancement of the operational definition of the test, along with a more detailed methodology for the procedures, would see an improvement in the reliability scores.

11.5.7 Hip extension test

Dynamic instability is a lack of control of the stabilisation responses required to maintain postural stability in a way that ensures maximum protection of the segments, and thus injurious movements are possible (138). This is seen in those with chronic conditions of the low back (138). The hip extension test is said to be a clinical sign of impaired motor control in the lumbar spine: the test itself involves an observed hip extension to the rear whilst lying prone. The authors investigated the inter-examiner reliability of this test for identifying dynamic instability. There were eighteen participants and two assessors included, one with less than a year of clinical experience and one with more than a decade. The Kappa agreement score was 0.72 and 0.76 for the right and left leg respectively. This is an encouraging result given the test's apparently high frequency of use within the clinical environment. The authors attribute these high levels of agreement to the simplistic nature of the test, the lack of equipment required, and the short amount of time needed to perform the test. Validity was not directly assessed here.

11.5.8 Hindfoot alignment assessment

Mal-alignment of the hindfoot has long been associated with foot and ankle injuries. Despite this, the literature has failed to adequately describe a reliable and valid method of hindfoot assessment (68). There is even a lack of adequate description as to what constitutes a varus or valgus pathology, or even the predictive ability of such an assessment (68). Both of these

issues need to be addressed as an essential prerequisite to further research in the area. Haight (68) aimed to determine the inter-rater reliability, intra-rater reliability, and test validity of the visual classification of hindfoot angle. In this study, examiners were required to measure the standing hindfoot angle of participants using a goniometer, then repeat their measurements by making a visual estimate of the same angle.

The intra-rater reliability of these measurements was found to be excellent (ICC of 0.88 – 0.90), and the inter-rater reliability was satisfactory (ICC 0.56 -0.65). In addition, within the experienced clinicians, visual assessments were equally as accurate as goniometry (ICC 0.64-0.95). On a less encouraging note, up to two degrees difference between trials was attributed to measurement error, and there were no normative values available with which to compare the measurements.

Frigg (150) developed this further by investigating the criterion based validity of visual assessments of hindfoot angle. Radiographical assessment (regarded as the gold standard of hindfoot assessment) of static hindfoot alignment was compared with assessments made visually during walking. The radiography assessment was based on the Saltzman-el-Khoury hindfoot alignment view method. Both of these methods report an angular value for the hindfoot.

In this case, only a 48% agreement level (no confidence interval reported) was found. This disappointing result was said by the authors to be partially attributed to the discrepancy between the bony alignment and what is viewed: the substantial amount of soft tissue present underneath the hindfoot may have produced a notably different observed visual angle to that seen using radiography. This explanation would appear to be logical, but remains to be tested.

Another issue in comparing hindfoot angle with radiography, which the authors did not discuss, is that the hindfoot angle is measured during weight bearing, whilst a radiography measurement is taken while lying prone. The differences in a joint while weight bearing,

compared to standing, may also have been a contributing factor in the low level of agreement. Finally, the use of a percentage agreement to report the results is not ideal; while it indicates scoring which is the same, it cannot adequately represent the proximity of scoring, which is also relevant in investigations of agreement (151, 152).

11.5.9 Proprioceptive ability or balance control: The single leg balance test

Using the theoretical basis of the kinetic chain (as discussed further in Chapters 6 and 7), injuries to the lower limb can be linked to the ability of the core muscles (those which surround the abdomen and spine, defined as the centre of postural control) to support and control the movements which occur distally (153). The core muscles are the foundation for the dynamic control of the torso, which in turn allows production, transfer and control of forces and movements within distal segments (153). The importance of the core has been highlighted in studies which have linked core muscle fatigue to subsequent hamstring injuries (154), and other studies which have linked proprioceptive ability of the core muscles and balance control to future injuries (100, 153-155). Three out of four of these studies were not eligible for discussion here as their measurement device did not fit the inclusion criteria: they required software packages and/or specialist equipment.

Trojian (153) studied 230 male and female collegiate athletes, all of whom were tested for balance control using one of the most widely recognised clinical tests of single leg balance assessment. Both inter-rater reliability (two assessors, forty participants) and test validity (230 participants, one assessor) of the single leg balance test were assessed in this study. It was found that there was a significant association between a failed test (inability to maintain balance with eyes closed for ten seconds) and subsequent injury (adjusted p-value <0.05). The inter-rater reliability of the single leg balance assessment produced a kappa score of 0.898, which was well above the authors' specified minimal acceptable value of $K = 0.6$.

The high level of agreement between raters here is thought to be due to the highly simplistic yet specific nature of the single leg balance test used.

11.5.10Sport specific tests

The use of battery-style functional tests for injury risk assessment is becoming increasingly popular (67). There are currently a large number of pre-season and functional tests available to the clinician (14, 67). These screening protocols use several tests in conjunction, which addresses an apparent limitation of the single screening test: an individual test cannot alone provide enough evidence to detect a 'functional abnormality', hoped to be predictive – in isolation – of overuse injury (67).

Pre-season screening is an injury counter measure utilised within many team-sports clubs by those in the physiotherapy and sports medicine fields. It is increasingly a requirement for participation in many sports at a professional level (67): for example, in New Zealand Triathlon each recognised athlete must complete pre-season injury screening annually (personal experience). Gabbe (67) identified what they termed 'five of the most common musculoskeletal pre-season tests': the modified Thomas test, the passive straight leg raise, active knee extension test, sit and reach test, and the active slump test. Gabbe (67) noted that pre-season assessments must be simple to perform, reliable, require minimal and inexpensive equipment, and be capable as a field test. They also must be highly standardised to allow comparisons across time and between clubs and sports, as comparisons are imperative within individuals over a season, and between athletes in different clubs.

Intra and inter-rater reliability assessments (fifteen participants, two raters, two repeated tests, one week separation) were performed on these tests collectively by comparison of scoring assigned by the individual raters. Results showed an inter-rater reliability of ICC 0.88-0.97, and intra-rater reliability of ICC 0.63-0.99. All of the ICC

assessments made were above the authors' pre-defined acceptable level of ICC = 0.60. Those authors attribute this finding to the simplicity and high level of standardisation of these tests.

11.6 Discussion

Table 11-1 presents the all findings from the included studies. It is apparent that the literature is most lacking in the area of validity, and the most commonly investigated psychometric property is intra-rater reliability. This is problematic, as different psychometric properties are all of equal importance to the practitioner. The reason for the lack of validity investigations, and the favouring of intra-rater reliability investigations is unknown and not discussed in the literature. It is speculated that the relative simplicity and convenience of an intra-rater reliability study has influenced this choice.

Seven of nine assessments of intra-rater reliability investigations produced results which met a clinically acceptable level, as chosen by the respective authors. The scoring ranged from very low (Gillet test, Kappa <0.08) to almost perfect (Sport specific test, ICC = 0.99). These findings are highly variable, which is to be expected given the wide range of functional tests included in the review. It is considered to be a positive finding that the majority of studies included met the clinically acceptable minimum level, which is also true in the results of the inter-rater reliability assessments (four out of seven achieving an acceptable agreement level). The inter-rater reliability scores had a similar range to the intra-rater reliability scores, also from very low (Gillet test, Kappa <0.08) to almost perfect (Sport specific test, ICC = 0.99).

It is interesting to note that of the studies which investigated both inter and intra-rater reliability, four out of five produced highly similar results across both reliability areas. For example, the scores for intra and inter-rater reliability of the Multi-hop test were both ICC>0.9, and for the Gillet test these were both Kappa <0.08. The reason for this is unknown, but it is suggested that for those tests which consistently scored well, the high level of test

standardisation and comprehensive nature of the scoring systems allowed for little error in assessment. For those which consistently scored below an acceptable level, a lack of test familiarity and perhaps ambiguity in the scoring system produced the variability present in the assessments. Both of these suggestions are also offered by the authors of included studies (67, 68, 153).

Three of the eleven included studies investigated criterion based validity, two of which investigated the hindfoot alignment assessment. Again, the results of these tests were highly variable. The hindfoot alignment view compared to radiography scored a 'low' 48% agreement level (150), HJAT comparisons were 'moderate' (147), and interestingly, the hindfoot alignment view compared to goniometry produced a score which was 'almost perfect' (ICC0.64 to 0.95) (68). Given the scarcity of the research on criterion based validity, and the variability of the findings, there cannot be any firm conclusions drawn from the collated evidence. While this is mainly due to a lack of available studies, it does not change the fact that practitioners do not have access to this information. The same issue is present in all areas of validity, were three studies investigated test validity and only two studies performed the comparison of construct validity. On a more positive note, the results of the test and construct validity investigations indicated that their respective assessments were 'supported' by the findings. In both instances of construct validity investigations, this was derived from a significant difference in the mean scores between injured and non-injured groups.

11.6.1 Limitations of included studies

Aside from the limitations already discussed, there were several other limitations identified among included studies. With the exception of Finnoff (145), none of the included studies were powered. As will be discussed in Chapter 7, it is important to ensure that a study has adequate power, because with the results of any scientific investigation, it is possible that an

association seen may have occurred due to chance. It is also possible that an association may be missed due to chance. Study numbers should therefore be big enough to minimise the occurrence of either.

Another limitation identified was that studies did not provide discussion on the potential for a learning effect of participants. This would not be so relevant to a test of reliability: which is essentially a test of how consistently a test can be graded, but could have an effect on investigations of validity. Studies only reported the number of repeated trials included as part of their methodologies, which ranged from twice (137) to six times (136), and as such, no conclusions can be drawn on the potential for a learning effect.

Finally, in all included studies, the experience of clinicians ranged from one (138) to ten years (138). Although this would undoubtedly have improved the reliability scores observed, it does mean that the results may not be applicable to novice examiners. This lack of knowledge of performance among novice practitioners represents a limitation of current research.

11.7 Conclusions

When considering the findings of previous investigations together, there are several points which create concern. Firstly, when investigating the reliability and validity of an individual test there needs to be some reference or comparison procedure. Unfortunately, there is a distinct lack of generally accepted diagnostic criteria on which to base the tests, and no gold standard methods of comparison for most of the clinical tests outlined. This makes assessments of validity particularly hard.

Secondly, those tests which can be compared to a 'gold standard' reference typically use radiographic measures. However, because radiographic assessments have excellent reliability, does not mean that they are clinically valid tests as we still do not know whether they are providing an indicator of function, and thus the risk of injury (54). This would be

especially true when considering the biomechanical properties of the hindfoot angle, as even though the hindfoot may appear ‘normal’ radiologically, the soft tissue padding underneath the calcaneus could alter function.

Finally, clinical screening protocols in general are based on the assumption that a biomechanical system which appears ‘normal’ leads to a low risk of injury: symmetry of function and absence of symptoms are closely related. There is some evidence to support this, yet it does not explain the high incidence of asymptomatic persons with apparent biomechanical abnormalities. Until a causal relationship between overuse injuries and asymmetries of form is established, such tests are at best suggestive (156).

This notwithstanding, there could be a plausible explanation for such reported positive findings in asymptomatic persons, which appears to imply a lack of specificity (where specificity is defined as the accuracy of the test in identifying true ‘positives’ (157)). In one study, between 16 and 33% of asymptomatic individuals were found to be classified as having a dysfunction (137). The positive test, or identified dysfunction in asymptomatic patients, is not necessarily a ‘false positive’, as it is possible that the subject may just not have experienced any symptoms yet. This is a purely speculative explanation at this stage, and requires further research. All clinical examinations need to be assessed for sensitivity and specificity, as this is an essential pre-requisite to understanding the information derived from the test.

Variable findings between studies which investigate the same clinical tests, can be partially explained by variation in use of terms. For example, the positions referred to as ‘neutral’ or ‘relaxed’ may not be the same between trials. In addition, the optimal position to measure such alignments is controversial. For example, Buck (158) states that a valgus position is an ideal point of reference when assessing the foot, while Frigg (150) states that

varus is optimal. The lack of consensus on what constitutes these postures needs to be addressed in future studies (159).

11.8 Summary

The main finding from this review is that the popularity of clinical tests is at variance with what is currently known regarding their psychometric properties, and in particular, their validity. This presents problems for the practitioner at every stage of the injury management system. As these tests are typically used as injury risk assessments, if such assessments cannot be made accurately, the correct identification of at-risk persons, and optimal selection of appropriate interventions cannot be achieved or their outcomes assessed.

Investigations of assessment validity were clearly lacking. In the case of criterion based validity, it is suggested that this may be partially attributed to the absence of a universally acceptable 'gold standard' reference. This represents a considerable challenge for future research in this area, and it is difficult to address this issue satisfactorily. For this reason, it is suggested that future research concentrates on the investigation of construct validity when investigating psychometric properties of visual assessments of function for injury risk: queries such as whether the test is a measure of what it claims to be (141). Such research would focus on ascertaining whether the visual assessment of function correlates with risk of injury over a specified period of time to establish whether there was a significant difference in injury occurrence based on the outcome of the visual assessment. Investigations of this nature are much more relevant to the practitioner too, as they are essentially a means of justifying the continued use of the test as an indicator of injury risk.

On a more positive note, investigations of intra- and inter-rater reliability were encouraging, with the majority of assessments reaching a clinically acceptable minimum level. Further to this, based upon the findings of current literature and the flaws identified, it is possible to implement a much better informed development process for future screening

tools. Function-based testing involving multiple assessment tools/components has become increasingly popular with practitioners and researchers alike. These are based on the assumption that identifiable biomechanical deficits in functional movement patterns have the potential to limit performance and make athletes susceptible to injury (160). According to the research presented here, this may represent a more robust approach for clinicians into the future, due to the increased reliability and validity of assessments apparent when multiple tests are used in conjunction. Based on these findings, no one test has supporting evidence regarding its psychometric properties. Multiple tests should be better predictors, but research in this area is lacking, as discussed here.

Clearly more research is needed to validate and confirm the reliability of other commonly used musculoskeletal tests. In particular, there were only three studies identified, which met the inclusion criteria, which attempted to *quantify* an aspect of validity of a clinical test (Munro et al (136), Frigg et al (150), and Haight et al (68), all investigations were of criterion based validity). This issue would be highly problematic for a practitioner who is seeking to make informed choices about their use of clinical tests, and clearly more research is warranted within the area of psychometric testing of functional assessments of overuse injury risk.

11.9 In the next Chapter

The following Chapter presents a systematic review of orthoses use as an injury prevention strategy within the military. It is a unique piece of research in that, as far as our knowledge goes, this specific investigation has never been performed before, and it was essential in answering one of the research questions for the current thesis:

What is the current evidence of clinical effectiveness of orthoses in reducing target injuries in the population of interest- the military?

The systematic review includes a meta-analysis and focussed discussion on the topic.

12 Chapter 4. Orthoses as an injury prevention strategy for lower limb overuse injuries in the military: A systematic review and meta-analysis of effectiveness

12.1 Overview of Chapter

The previous Chapter concerned visually-based functional screening assessments currently available to the practitioner. The discussion focussed on screening tests, and reviewed investigations of psychometric properties of such tests, including salient findings from these studies.

The current Chapter presents a systematic review, including meta-analyses, focussing on the effectiveness of orthoses use as an injury prevention strategy for lower limb overuse injuries in the military. It specifically addresses one of the main research questions of the current thesis (refer to Chapter 1):

What is the current evidence of clinical effectiveness of orthoses in reducing target injuries in the population of interest- the military?

12.2 Introduction

Because of the high prevalence of lower body overuse injuries in the military (80) (refer to Chapter 2), there has been considerable research exploring various preventive interventions (35, 97, 107). Aetiology of such injuries have been associated with repetitive and cumulative adverse tissue loading, leading to: stress fractures within the bones of the feet, enthesopathies involving tendons and fascia in the feet, fascial and muscular compartment syndromes in the lower leg, tibial stress reaction syndromes, tibial stress fractures, as well as contractile, ligamentous, and intra-articular injuries to the knee, hip and lumbo-pelvic regions (123, 125, 130, 161).

Footwear inserted orthotics are frequently used in the military as both a primary and secondary intervention. As a primary intervention, these are part of a preventive strategy to reduce injury risk, and as a secondary intervention for those diagnosed with the disorders and conditions indicated above (123, 161). Given that the risk of further injury can increase by up to 400% once an individual has sustained an overuse injury of the lower limb (109, 121, 162), with substantial associated treatment and management costs (123), effective injury prevention represents the most attractive option.

Due to the frequency of these injuries in the military (refer to Chapter 2) various research groups have investigated the effectiveness of the use of orthoses as an injury prevention strategy, which justifies the current systematic review focussing on research in military populations. Another important driver for the current review is the apparently widespread practise and use of orthoses in the military (New Zealand and overseas, refer to Chapter 2). This represents the first such review, as previous reviews have included research carried out on various populations: a search of the published literature identified several systematic reviews which investigated the use of orthoses as an injury prevention strategy, collating evidence from both athletic and military populations (123, 125, 130, 161). These reviews generally concluded there was insufficient high quality evidence to either support or refute the clinical effectiveness of such orthoses use in military and athletic populations.

12.2.1 Objectives

This review therefore aims to determine the current evidence of effectiveness to allow informed decisions to be made on the use of orthoses as a potentially affordable and feasible intervention strategy for the military. The main aim of this review was to systematically review the evidence (including meta-analysis where appropriate) from prospective randomised controlled trials (RCTs) investigating the effectiveness of foot orthoses use for the primary prevention of lower limb overuse injuries within the military.

12.3 Methods

This was a systematic review of randomised controlled trials incorporating orthoses for primary injury prevention in the military.

12.3.1 Eligibility criteria

Based on current knowledge related to overuse injuries, four critical inclusion criteria were developed and applied to determine the most appropriate studies for this review, based upon PICO principles (i.e. population, intervention, comparison and outcome) (163). These are outlined and justified as below:

12.3.2 Population: Types of participants

This was defined by the criteria: "*Did the study use only military personnel as the participants?"*

As highlighted in Chapter 2, the unique nature of the military training environment dictates that it is appropriate to include only those studies which are completed utilising military personnel as the participant group. Carrying a rifle (thus altering the natural arm motion during gait (64)); and absolute load carried appear to be the only variables identified so far that consistently associate with overuse injuries in the military (refer to Chapter 2). Further, absolute load carried logically increases ground reaction forces, medio-lateral forces and time spent in stance, which are indicative of changes in gait (70). These factors are essentially military in nature and comparatively rare occurrences for other groups (e.g. sportspeople), thus, valid comparisons cannot necessarily be made between results obtained in military and non-military groups.

12.3.3 Intervention: Types of interventions

The second inclusion criterion was: "*Did the study solely focus on the use of orthoses as a primary preventative strategy?"*

The basis of this criterion is straightforward and is derived from the aim of the review. A primary preventative strategy is prescribed to individuals who are not currently, or 'recently', injured. It is important to make a distinction here because individuals who are currently suffering, or have recently suffered from lower limb overuse injury are up to 400% more likely to sustain a recurrence (109, 121, 162). True risk depends strongly on the presence or absence of an initial injury and is worse for stress fractures. The investigation of orthoses use as a treatment and for subsequent prevention will be skewed for this effect.

12.3.4 Comparison: Types of comparisons

The inclusion criteria concerning the studies comparison was defined by: "*Did the study incorporate a parallel control group, without the use of a placebo?"*

For the purposes of this review, a prospective study with a parallel control group, without the inclusion of a placebo, based on random allocation, was chosen as the required comparison. A parallel control group was considered necessary given the dynamic and varied nature of military training and its environment, any comparisons to historical or previous data would likely be invalid. Seasonal differences, for example, can alter the training terrains and potentially affect the rate of injury occurrence, arguably reducing the validity of such comparisons.

The exclusion of studies which incorporated the prescription of a placebo insert to those in the control group is justified by a recent review of the orthoses literature (129). Any material which makes contact with the skin has the potential to alter function through the stimulation of cutaneous receptors, which has been illustrated in research on strapping tape (101). Thus, as discussed in Baxter et al (129), there can arguably be no such entity as a placebo shoe insert. Because of this, studies which provided the parallel control group with a placebo insert were excluded.

12.3.5 Outcomes: Types of outcome measures

The final criterion was: *"Were indicated injuries monitored with follow up of participants prospectively for a minimum period of three months?"*

Injuries to bone (stress reactions and stress fractures) can take eight (164) to sixteen weeks to correctly diagnose (165). Therefore, any research taking less than three months may miss such bone injuries that take considerable time to identify (165). Although previous reviews have used a shorter time frame, a three month minimum timeframe was deemed most appropriate. The injuries of interest were any type of overuse injury to the lower limb, monitored within a study which met the inclusion criteria.

12.3.6 Information sources

The following databases were electronically searched during a three month period from May to July 2011, from inception to present: CINAHL, Cochrane, Medline, Pubmed, Scopus and Web of Science as well as a manual search of the reference lists of selected relevant articles. Only articles which had been written in, or translated to, English were selected. Further, articles had to have undergone a peer review process to be included. An example of the search strategy used is outlined below.

12.3.7 Search

Database: Ovid MEDLINE(R) 1948 to Present with Daily Update

Search Strategy:

---

1 exp Shoes/ or exp Foot/ or exp Orthotic Devices/ (45497)

2 exp Military Personnel/ (19641)

3 exp "Wounds and Injuries"/ (608840)

4 1 and 2 and 3 (65)

****

12.4 Data collection and analysis

12.4.1 Study selection

The primary investigator (MLB) completed the search in its entirety following the process outlined, and completed the search of abstracts. Studies to be retrieved in full text were identified. The second reviewer (DCR) then, independently, repeated this search strategy in its entirety following the same process, and also identified full text articles meeting the criteria. A third reviewer was consulted for majority opinion when there was a lack of consensus between the two primary reviewers. The justification for inclusion or exclusion of any study, at all three levels (title search, abstract search, full text) was based solely on the four criteria outlined.

12.4.2 Data collection process

Both MLB and DCR independently extracted data from the included trials. Disagreements were clarified through discussion and consensus. The results gained from each of the retrieved studies were pooled and a meta-analysis was performed. The comparison made was injury occurrence between orthoses and control groups in the military for each specific injury.

Data items

Data pertaining to the participants' injury rates were extracted, as well as any items relevant for quality assessment (as outlined below).

12.4.3 Risk of bias in individual studies

All full text articles retrieved and selected for inclusion were assessed for risk of bias. The assessment of quality was performed independently by two assessors (MLB, DCR) using the criteria outlined in Table 12-1, which is an adapted version of the PEDRO scale; item six was removed as it was not relevant for this investigation *("Item 6: there was blinding of therapists administering the therapy")* (166). Studies were rated for bias as either zero or one. Zero score was assigned where there was insufficient information or if there was a high risk of bias. One was assigned where there was a satisfactory measure in place to reduce bias. Any discrepancies between authors' assessment were clarified through discussion and consensus. A kappa test was performed to assess the initial agreement between the authors.

Table 12-1: Numbered items of quality assessment as per the PEDRO scale (166).

Number allocated	Description of quality (1 = yes, 0 = no or impossible to tell)
1	The eligibility criteria were specified explicitly
2	Participants assignment to groups was random
3	Allocation of intervention was concealed from the participants
4	The groups were similar at baseline for all of the most important indicators or known risk factors
5	The participants were blinded to the expected outcomes
6	Those who were assessing the outcome were blinded to the participants intervention status
7	Measures of at least one key outcome were reported from at least 85% of the original group
8	At least one intention to treat analysis was performed, or all members who received treatment were included in the outcome measures
9	The results from a between group statistical comparison was reported
10	The study provides measures of variability for at least one key outcome, for example, number of outcomes in each group

12.4.4 Summary measures

Forest plots pertaining to specific injuries were created using R[c] statistical software (cran.r-project.org), and relative risk values were calculated for each individual injury.

12.4.5 Synthesis of results

All data relating to specific injuries were pooled and included in each forest plot within the meta-analysis. The statistical analysis was completed based on injury occurrences. An injury, as defined by all included studies, was one that was diagnosed by a military appointed medical practitioner, which required a minimum of one day off duty. Heterogeneity of studies will be calculated within the meta-analysis performed for each injury individually, (chi-squared tests shown in Figures 4 – 13).

12.5 Results

12.5.1 Study selection

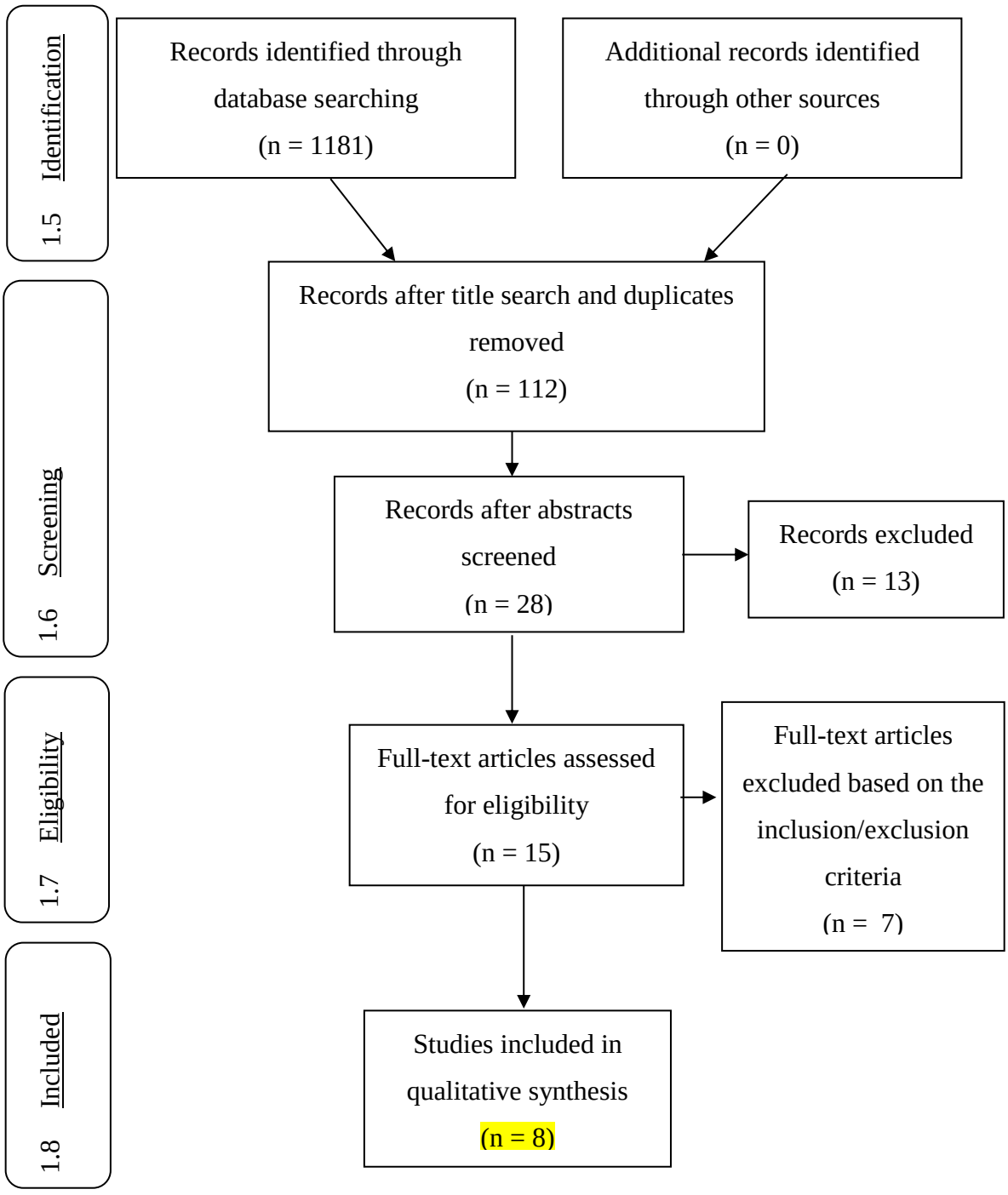


Figure 3: Diagram showing number of retained studies during the search process.

Number of retrieved and retained studies are shown in brackets.

Table 12-2: Characteristics of the eight studies included in this review

Author and Date	Investigation	Participants	Interventions	Outcomes	Results and conclusions
Finestone 1999 (167)	The effect of issuing custom made biomechanical insoles on the incidence of stress fractures was tested. Two insole groups and one control group.	Israeli infantry basic training recruits n=323	Fourteen weeks of either orthoses or control groups	Stress fracture and specific location incidence rates	No statistically significant reduction in injuries. There was no justification for the prescription of the biomechanical orthoses
Gardner 1988 (168)	A controlled trial determining the usefulness of a shock absorbing insole in prevention of stress fractures and stress reactions of the lower extremities: An intervention group and a control.	3025 US marine recruits	Either an insole or no insole for twelve weeks of basic training	Stress fractures and 'other injuries' were recorded	No significant evidence for the mass prescription of orthotics for stress fracture prevention. Some specific injuries may be responsive.
Sherman 1996 (169)	Prevention of lower limb pain using shock absorbing insoles; either given orthoses, control, or purchased on their own.	1132 male basic trainees US army	Either shock absorbing insole, motion controlling or no insole	Specific injury rates of the lower limb and lower back were recorded for incidence rates. (duration unspecified)	All injuries were reduced for the intervention group, not all were statistically significant (3/7).
Larsen 2002 (63)	Custom made insoles for the prevention of lower back and lower limb injury; intervention and control groups	146 Danish military basic training recruits	Either a custom made biomechanical insole or no intervention.	Specific injury incidence rates were collected over three month period.	No significant differences in injuries, in the worst-case analysis. Certain conditions may be preventable with orthotic prescription.
Milgrom 1985 (170)	Shock absorbing orthoses to reduce stress fractures in military recruits; orthoses group and control group	295 male infantry recruits Israeli army	Shock absorbing insoles or no intervention for fourteen weeks of basic training	Incidence rates of femoral, tibial and metatarsal stress fractures were recorded.	Significant reductions were observed in the cases of tibial only.
Milgrom 2005 (171)	Biomechanical orthoses for the prevention of weight bearing induced lower back pain: one intervention group and one control.	404 Israeli army infantry recruits	Either a foot orthotic or no intervention for fourteen weeks of basic training	Incidence rates of back pain were recorded.	Use of orthotics is not recommended to prevent lower back injuries.
Mattila 2011 (172)	Shock absorbing, customised orthotics for the prevention of lower limb overuse injury. Orthotic vs control at a 1:2 participant ratio.	220 Military conscripts Finnish Defence Force	Customised orthotic or no intervention for the first six months of service	Incidence rates of any lower limb overuse injury, which required at least one day off work.	Use of orthotics is not recommended to prevent lower back injuries.
Simkin 1989 (173)	Biomechanical orthoses for the prevention of stress fractures: one intervention group and one control	256 Israeli army infantry recruits	Either a foot orthoses or no intervention for fourteen weeks of basic training	Incidence rates of femoral, tibial and metatarsals were recorded	Orthotics reduced stress fracture incidence in the case of tibial and metatarsal only.

12.5.2 Study characteristics

Table 12-2 outlines the characteristics of included studies. Eight studies met the criteria for inclusion in this review (63, 167-173). Seven full text articles which were excluded as they did not meet the inclusion criteria. Six of these did not provide evidence of sufficient follow-up (which was a minimum of three months), and the final full text article contained summary data which was already present in another included study.

12.5.3 Risk of bias in included studies

Included studies were ranked on a bias score from zero to ten. A score from zero to five indicated a study of low quality while a study score from six to ten was considered to be high quality (166). Five studies were ranked as being of low quality, having high bias (167, 169-171, 173), while three were ranked as being high quality, and low bias (63, 168, 172). Scores ranged from three to eight. The risk of bias in included studies was assessed using the PEDRO 'Assessment of bias' table. Results for risk of bias assessment are presented in Table 12-3.

Overall, based on the GRADE Working Group (PEDRO) level of evidence scale (166), this corresponds to a 'low quality of evidence'. The studies with the lowest quality assessment scores were Finestone et al (167), which had a score of three out of ten, and Simkin et al (173) which had four out of ten. The studies with the highest quality assessment scores were Gardner et al (168) (seven out of ten) and Larsen et al (63) (eight out of ten). Many aspects of methodological quality were impossible to assess due to there being a lack of relevant information in the article (and thus in keeping with PEDRO scoring guidelines these cases were also marked as zero). The lowest rated studies of Simkin et al (173) and Finestone et al (167) also had the greatest lack of clarity in their study designs (both had

insufficient information to assess risk, and corresponding score of zero, in four out of ten aspects).

Some of the most notable identified weaknesses included: Finestone et al (167) only included those who were 'completers' in their analysis, which was only around 50% of the original participant group; attrition rates were therefore nearly 50%. Similarly, Simkin et al (173) and Milgrom et al (170) did not report the results or outcomes from over thirty of the original participant group due to their discontinued use of orthoses.

Table 12-3: Assessment of quality for each of the studies included in the systematic review.

Author and Year	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Total /10
Finestone 1999	0	1	0	0	0	0	0	0	1	1	3
Gardner 1988	1	1	1	0	0	1	1	1	0	1	7
Larsen 2002	1	1	1	0	0	1	1	1	1	1	8
Milgrom 1985	1	1	0	0	0	0	1	0	1	1	5
Milgrom 2005	1	1	0	0	0	0	1	1	1	0	5
Mattila 2011	1	1	0	0	0	1	1	1	0	1	6
Sherman 1996	0	1	0	1	0	1	1	0	0	1	5
Simkin 1989	0	1	0	0	0	0	1	0	1	1	4

12.5.4 Between assessor agreement

Kappa coefficient, used for assessing inter-rater reliability between MLB and DCR, was described as 'good' (Cohen's Kappa = 0.654) (151). Given the low level of clarity in many of the studies, this was considered a very satisfactory result.

12.5.5 Meta-analysis and synthesis of results

The broad definition of the intervention investigated can be referred to as: "*orthoses versus control for the reduction of lower limb overuse injury"*. It is important to acknowledge that

there are two different types of insoles available, proposed to be 'shock absorbing' (169) and 'motion controlling (biomechanical)' (171). Shock absorbing insoles are made from a visco-elastic polymer and were used in four of the eight studies (168-170, 172). Motion controlling insoles are made typically from multi-layered memory foams, and were used in the remaining studies (4/8) (63, 167, 171, 173). Due to the lack of data for each orthoses design category, the results from all studies were pooled together for meta-analysis; however the potential issues with this need to be acknowledged.

Forest plots obtained from injury comparisons were as follows:

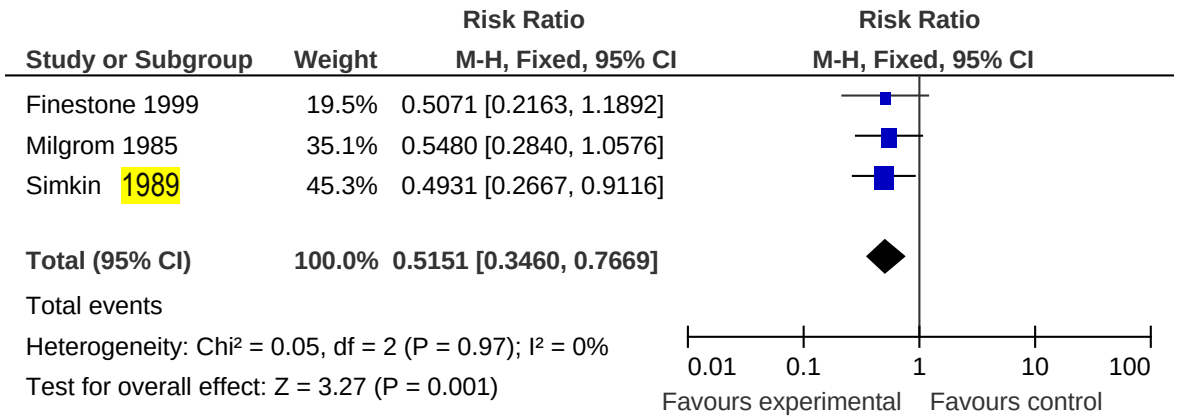


Figure 4: Forest Plot for Femoral Stress Fracture, Data obtained from three of the eight included studies. Data represents incidence of femoral stress fractures between experimental (orthoses) and control (no orthoses) groups.

The data above show a trend which favours the experimental (orthoses) groups. The pooled results show a significant effect of injury reduction in the orthoses group for femoral stress fractures. The chi-squared test results indicate that there are no significant differences between expected (i.e. favouring the experimental group) and observed data. The overall affect shown by the z-value indicates that there is a significant difference between control (no orthoses) and experimental (orthoses) groups.

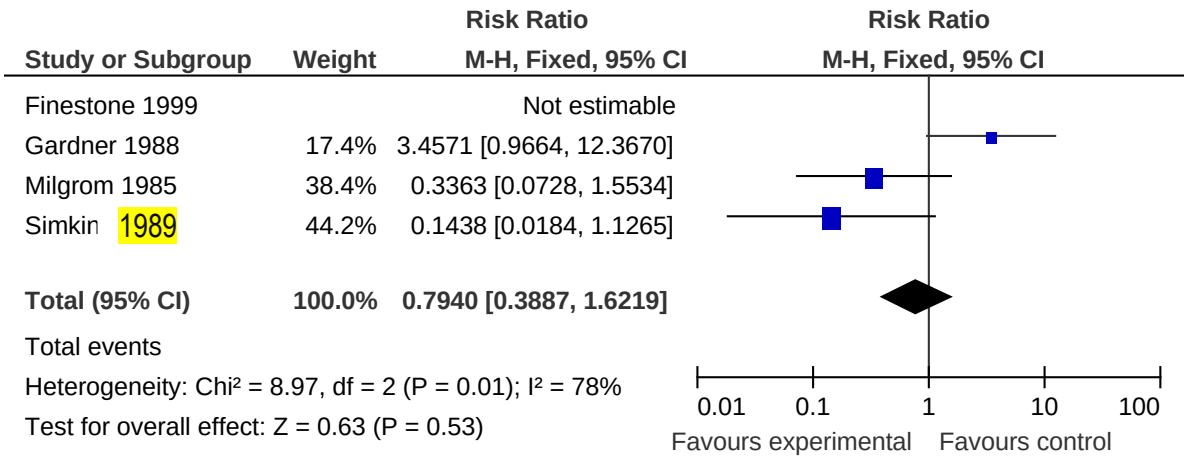


Figure 5: Forest Plot for Stress Fractures of the Foot. Data obtained from four of the eight included studies. Data represents incidence of stress fractures of the foot between experimental (orthoses) and control (no orthoses) groups.

There is no definite trend shown in Figure 5. There were no instances of relevant injury in either the experimental or control groups for the study performed by Finestone et al (167). Results of all the data sets include 1.0 and thus are not significant; no effect is seen. The chi-squared test shows that there is a significant difference between the expected and observed data: the orthoses intervention group was not significantly different from the control group. The z-value for the overall effect indicates that there are no significant findings.

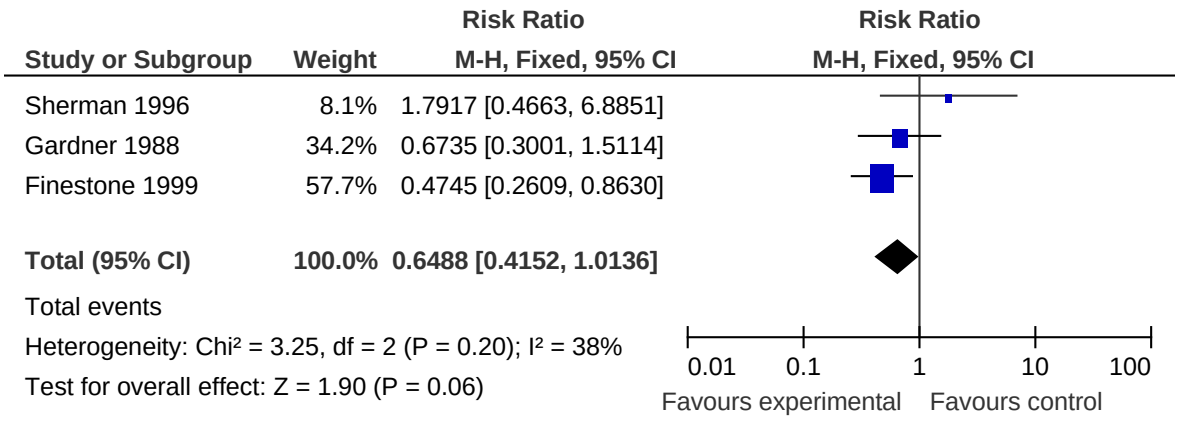


Figure 6: Forest Plot of General Stress Fractures (includes any lower limb stress fracture, with the exceptions of the tibia or femur). Data obtained from three out of the eight included studies. Data represents incidence of stress fractures referred to as 'general' between orthoses and control groups.

In all the studies apart from Finestone et al (167) the 95% confidence interval included 1.0 and therefore no effect was seen. The pooled data also included 1.0. The expected results were not observed ($chi^2 = 3.25$) and there were no significant findings ($Z = 1.90$, $P = 0.06$).

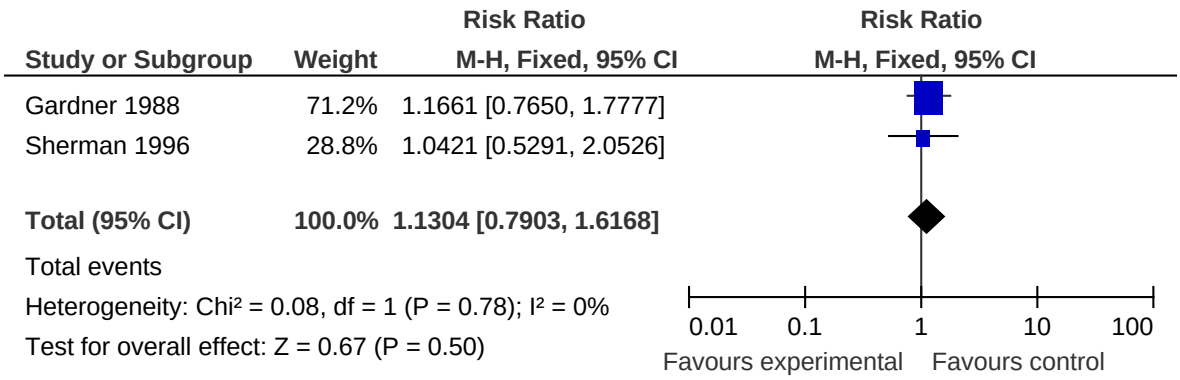


Figure 7: Forest Plot of Overuse Injuries of the Foot. Data obtained from two out of eight included studies. Data represents incidence of overuse injuries of the foot between orthoses and control groups.

All data are centralised around 1.0 and thus the results are equivocal. There were no significant findings observed (Z=0.67, P = 0.50).

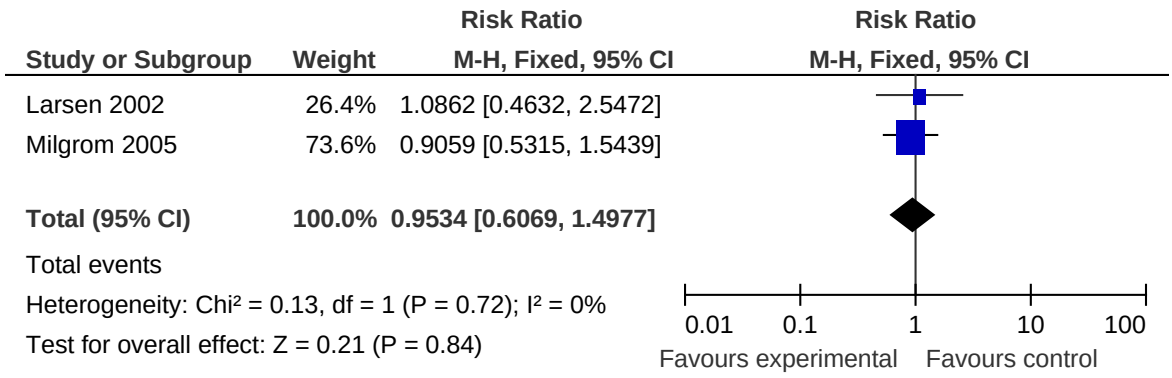


Figure 8: Forest Plot of Overuse Injuries of the Back. Data obtained from two out of eight included studies. Data represents incidence of overuse injuries of the back between orthoses and control groups.

All data are centralised around 1.0 and thus the results are equivocal. There were no significant findings observed (Z=0.21, P = 0.84).

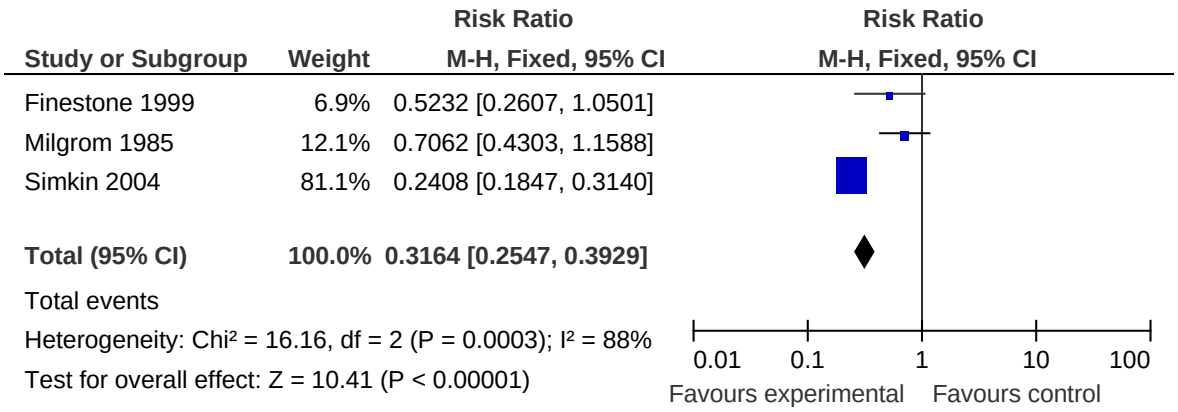


Figure 9: Forest Plot of Stress Fractures of the Tibia. Data obtained from three out of eight included studies. Data represents incidence of stress fractures of the tibia between orthoses and control groups.

All data favour an effect from using orthoses intervention, although two of the three data sets used have a confidence interval that includes 1.0. The pooled data indicate a significant effect of orthoses use on reducing the incidence of stress fractures of the tibia ($Z= 10.41$, $P<0.00001$). The chi^2 test indicates that the expected results were observed, and the overall effect size was significant.

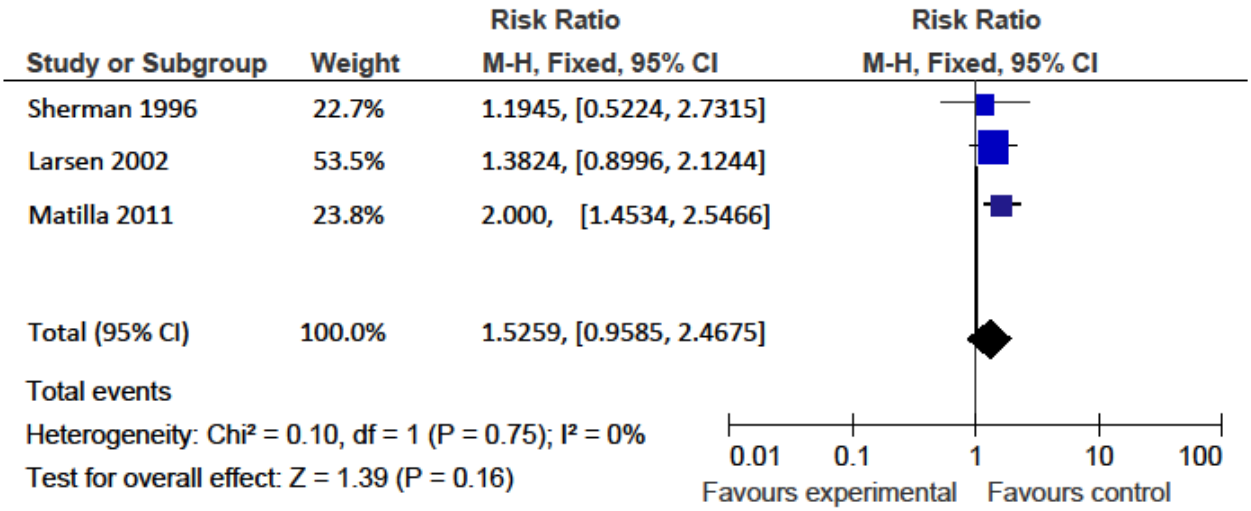


Figure 10: Forest Plot of General Overuse Injury of the Knee. Data obtained from three out of eight included studies. Data represents incidence of general overuse injury of the knee between orthoses and control groups.

All data are centralised around 1.0 and thus the results are equivocal. There were no significant findings observed (Z=1.39, P=0.16).

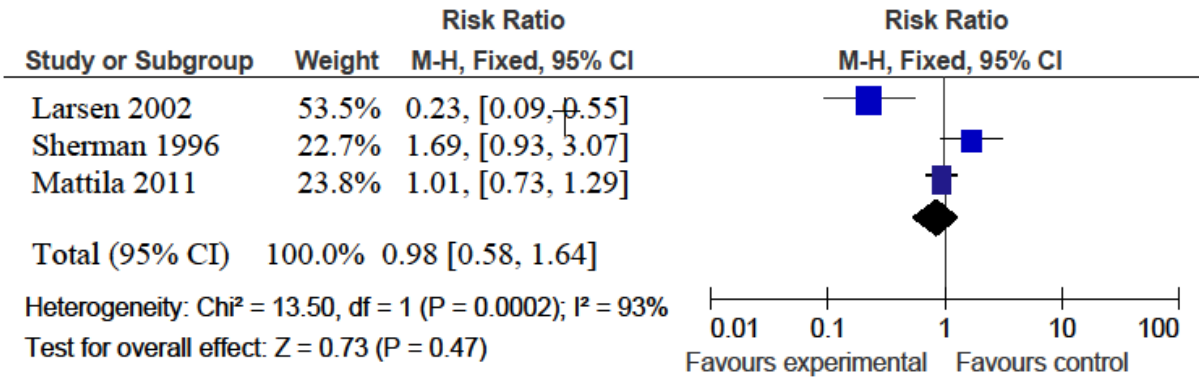


Figure 11: Forest Plot of Medial Tibial Stress Syndrome. Data obtained from three out of eight included studies. Data represents incidence of medial tibial stress syndrome between orthoses and control groups.

There is no definite trend shown in Figure 11. The results from one of the data sets and the pooled results include 1.0 and thus are not significant: no effect is seen. The chi-squared test shows that there is a significant difference between the expected and observed data. The z-value for the overall affect (0.73) indicates that there are no significant findings (P=0.47).

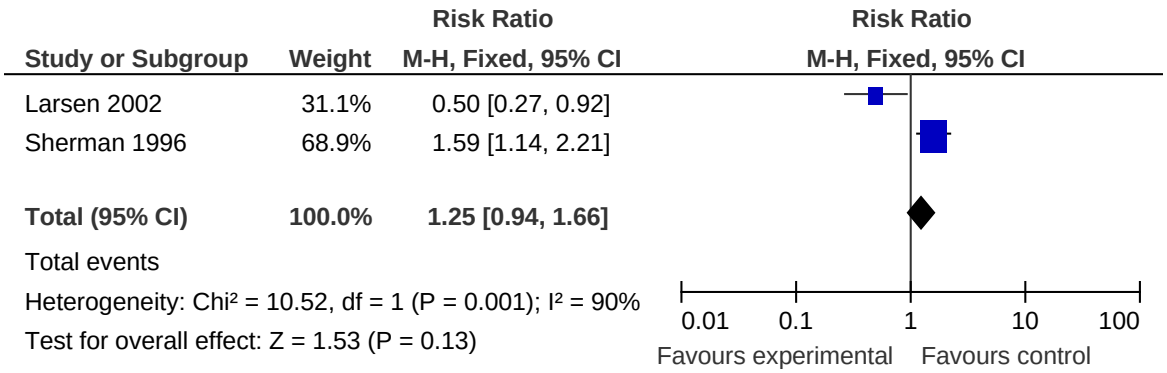


Figure 12: Forest Plot of Sprains and Strains. Data obtained from two of the eight included studies. Data represents incidence of sprains and strains between orthoses and control groups.

There is no definite trend shown in the results. Interestingly, one of the data sets significantly favours the experimental group, while the other data set significantly favours the control group. The pooled data includes 1.0 and therefore is not significant. The chi-squared test shows that there is a significant difference between the expected and observed data. The z-value for the overall affect indicates that there are no significant findings.

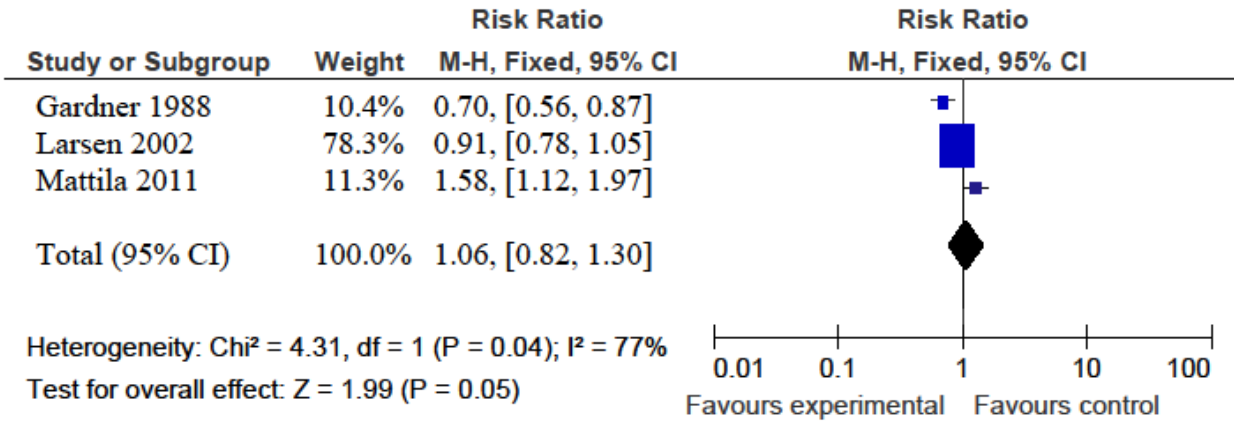


Figure 13: Forest Plot of Other Overuse Injuries. Data obtained from three of the eight included studies. Data represents incidence of any other overuse injury (unspecified) between orthoses and control groups.

All data is approximately centralised around 1.0 and thus the results are equivocal. Two of the three data sets and the pooled data set includes 1.0. There were no significant differences observed between groups (Z=1.99, P=0.05).

12.6 Discussion

12.6.1 Summary of evidence

This review investigated the effectiveness of orthoses intervention as a means of preventing overuse injury among military personnel. In summary, the findings of this review and meta-analysis indicate that orthoses were favourable when compared to no intervention only in the cases of tibial and femoral stress fractures. For the remaining eight investigated injuries, relevant meta-analysis demonstrated no significant pooled effect. Due to the lack of data within the included studies, no statement can be made regarding the cost-effectiveness of this intervention. It is important to note that the individual injuries considered as part of this review were only represented in a limited number of included studies (in most cases three), and therefore the results lack robustness and generalizability in most cases.

There were apparently a large number of published articles available which directly related to the use of orthoses as an injury prevention strategy for the military. However, only eight studies were identified which met the inclusion criteria for this systematic review. Of these, five studies were ranked as being of low quality (167, 169-171, 173) and only three were ranked as being high quality (63, 168, 172). The overall assessment of quality level was 'very low'. There were several instances where no definitive conclusions on the quality of the study could be made, as there was insufficient information in the published account of the study. This was the primary flaw of the chosen group of studies, and confounded accurate assessment of the quality and risk of bias within the reviewed research. Lack of information was noted as a substantial contributing factor to the poor quality scores of the studies. These points highlight the need for further development of the field and more specifically, the urgent need for additional high quality studies.

12.6.2 Comparisons with other reviews

The current findings are similar to those of three similar systematic reviews which investigated the use of orthoses as an injury prevention-intervention (125, 130, 161). Two of these concluded that there was not enough evidence to make strong conclusions; however it was noted that the evidence indicated that orthoses did not inflict harm (130, 161). Findings from the third systematic review also did not provide compelling evidence for effectiveness, but did conclude that orthoses were the most promising of all intervention strategies (125). It is suggested that the current situation will continue until issues regarding methodological quality are addressed.

Another key issue, similar to that identified by Rome et al (161), related to the use of randomisation processes. In order for the control and intervention groups to be directly comparable, there needs to be a robust randomisation method for allocating participants to the intervention and control group (63, 167). Many studies which were included in this review

had questionable methodology with respect to randomisation (169-171, 173). In some cases, the method for assigning participants into groups was satisfactory (for example: picking a card out of a hat (167), or a box (63)), but in most cases it was impossible to make conclusions on the appropriateness of the randomisation due to lack of information, and therefore the inter-group comparability was questionable.

12.6.3 Effectiveness

The results of the meta-analyses show highly varied findings, which is interesting given the highly specific nature of the inclusion criteria. As a combined effect, the findings of this review indicate that the use of orthoses has no effect on overuse injury prevention for the military. Only for the injuries of 'femoral stress fractures' (Figure 4) and 'tibial stress fractures'(Figure 9) were significant results observed. However, although the results of the meta-analysis are generally weak, it seems that for some other investigated injuries orthoses *may* produce the desired effect; i.e. reduction in injuries, for 'general stress fractures' (Figure 6), 'other overuse injuries' (Figure 13) and possibly 'stress fractures of the foot' (Figure 5).

It was interesting to note the highly variable results of the meta-analyses. In some cases, for example Figure 12, there are clearly disparate findings from the three studies (63, 169, 172). Consideration of an overall effect is therefore difficult to interpret, as it simply represents a mid-way point between three studies with highly varied findings. While reasons are unclear, possible factors might include differences in the training programs and consequential differences in risk of overuse injury, or in the case of Figure 13, the use of different types of insole (one motion controlling, two shock absorbing).

12.6.4 Clinical relevance

Due to the poor methodological quality of the trials, no firm conclusions can be drawn on effectiveness, and this of course affects the clinical relevance of the results. Beyond this, in

several of the studies, large numbers of participants were not included in the reported results for various and often poorly defined reasons. In addition, there were high numbers of drop-outs in several of the studies: attrition was as high as 50% in some cases (167). This and the absence of intention-to-treat analyses limits any definitive conclusions. Thus, it is difficult to assess the cost-effectiveness of orthoses interventions using such an en-masse supply approach.

Speaking primarily at an injury reduction level, the evidence from this systematic review indicates that the provision of orthoses is more favourable in terms of injury prevention, to no intervention, for only some overuse injuries. Therefore, it is likely that they would not be cost-effective as an intervention. However, there is insufficient information supplied to make a definitive statement on the cost-effectiveness of this approach. Regardless of this, it is suggested that a targeted intervention of orthoses would potentially be more effective in reducing injury rates, and therefore be cost-effective, through the reduction of orthoses purchase costs. However, the absence of studies which utilised a targeted intervention of orthoses does not allow cost-effectiveness to be commented on further here.

12.6.5 Strengths and limitations of this review

The current review is limited in its conclusions given the poor quality of most of the included studies. In saying this, given the flaws in current methodologies, it has highlighted the importance of, and direction for, future research. It is disappointing that given the high number of relevant studies potentially available there were only eight which met the criteria for inclusion, and that of those seven all showed at least one area of high risk of bias. However, using the recommendations from this review, future research should provide more robust results.

The main strength of this review is that it only includes RCTs which utilised the target population (i.e. the military) as participants. This means that the findings are less susceptible to selection bias, and can be validly generalised to the target population.

A drawback of this review is that due to the lack of quality research available, there was a pooling of results from studies which used both shock absorbing and motion controlling insoles. This was not ideal, as it is known that the insoles are fabricated from different materials and therefore it can be speculated that they have different qualities and – potentially – effectiveness. Previous research comparing different types of insole, i.e. between shock absorbing and between material compositions, has concluded that different insoles can have different kinematic and kinetic effects (81). Unfortunately, there is insufficient literature to allow meta-analyses focussed on one specific type of insole. Thus, the orthoses potentially most applicable to reducing injury risk is as yet unclear.

A second drawback of this review was the lack of results available for meta-analysis. Of the eight studies included, three or fewer had specific results which could be pooled for each injury (refer to Figures 4-13). This decreases the reliability of the conclusions made for each injury, as a small number of studies included in any meta-analysis leaves the risk of any one study over-influencing the results. This effect is seen most clearly in Figure 8 and Figure 9 where one study accounts for around 80% of the overall result. Figure 9 is particularly noteworthy, where the size of one study has outweighed the equivocal findings of two other studies, and thus has over-influenced the pooled result.

12.6.6 Implications for research

It is important to note the highly variable results found from orthoses use for injury prevention. Clearly, this area is justified for further investigation, with more stringent methodologies. It is also suggested that further study should investigate whether orthoses

decrease *the severity* of injuries sustained, as all systematic reviews thus far have concentrated on rate or risk of injury.

Based on the findings of this review, it is recommended that future research be adequately powered and better designed to produce more robust conclusions. As alluded to in the discussion of clinical relevance, it is also strongly suggested that future research aims to identify screening strategies or risk factors in order to effectively determine whether orthoses should be supplied as a targeted intervention. Building on the discussion of risk factors, an area of research which is currently lacking is the knowledge of injury risk in different training scenarios; specifically, is it still unknown whether marching poses the same injurious risks as running when in full military kit. It may be the case that orthoses are more suitable as a preventative measure in only one of these modalities. Future research is needed in order to investigate this suggestion.

The general consensus of the current literature investigating the injury prevention capabilities of orthoses appears to be equivocal. This may be because no study identified how to apply this intervention based on individualised risk assessment. Therefore, it is speculated that current results could be based on whether the intervention group had more or less individuals who were predisposed. Although orthoses intervention could still be applied en-masse to the military, and therefore anyone who was predisposed would get the required intervention, the cost-effectiveness of this method is questionable and the sum effect may actually "wash out" the significance of the intervention. Individualised screening for functional and/or morphological injury risk factors may be the most important direction for future research, as exemplified in a recent study (1).

12.7 Conclusions

Based on the meta-analysis of pooled results and the moderate quality assessments, there can be no strong conclusions made on the effectiveness of the use of orthoses as a preventive

measure for overuse injury in the military. The findings from all of the included studies, of which five were ranked as being low quality, did not produce compelling evidence in favour of their use or non-use. Although the majority of results (seven out of ten overuse injuries) found in favour of orthoses use, the confidence intervals and relative risk values did not provide convincing results. Importantly, this finding was not limited to the studies which were identified as having the highest risks of bias. In addition, the high attrition rates and lack of intention to treat analyses within the included studies limits the robustness of these results. This is a flaw of current research that needs to be addressed in future studies.

The findings of this systematic review do not support the use of orthoses as a primary intervention for overuse injury prevention in the military. Therefore, based on these findings, the use of the orthoses as the standardised intervention to be used with FootFAST is not supported. Although it is not the purpose of the current thesis to identify an alternative, further research is needed in order to identify a suitable intervention. This topic is discussed further in Chapters 7 and 8.

12.8 In the next Chapter

The next Chapter outlines the background of the FootFAST protocol, and provides a justification for the choice of the three tests which comprise it. Following that, based on the findings from the literature reviews, a feasibility study was the necessary next step. This was to ensure the main study produced optimal results by ensuring the robustness of the methodology, and results from the feasibility study were used to inform the participant and assessor numbers required in order to produce sufficient power within the main study (80%).

13 Chapter 5. The FootFAST protocol

13.1 Overview of Chapter

The current Chapter outlines the FootFAST protocol in detail, then provides essential background information regarding the development of FootFAST and the choice of the three tests which comprise it. This leads on to the next Chapters, which describe the investigations into the psychometric properties of the FootFAST assessment. Although the FootFAST protocol was designed by Dr Charles Baycroft, the justification of the three tests presented in this Chapter was completed by the author (MLB).

13.2 Introduction

It was noted as far back as 1988 that a preventive, evaluative tool that could assess for risk of overuse injury in the lower limb would be useful for both patients and therapists alike (174). However, factors for overuse injury are complex and there is still relatively little known regarding the prediction of these injuries (175, 176). There have been a number of attempts to design screening tests for overuse injuries of the lower limb (137, 145, 149). The earliest assessments included testing of the spine, hips, knees, ankle, subtalar and midtarsal joints (174); however, there was limited success with this series of assessments. While it was noted that specific functional differences were associated with particular injuries, the relationship was complex, in that every individual with a measurable 'difference' did not necessarily have an injury; this is also true for many other clinical assessments (175, 176). This highlights the multi-factorial nature of injury risk. More recently, it was concluded from work with sporting athletes that injury was associated with a combination of many influential factors, including, but not limited to, the athlete's training program, footwear choice, and playing surface (177).

Because of the limited predictive ability of traditional screening tests for overuse injuries of the lower limb, these assessments have received little attention in the research literature, and many of the findings that are available are not encouraging (as discussed in the narrative review, Chapter 3). The use of such simple, 'single item' tests within a clinical setting is therefore difficult to justify. Given that risk factors associated with lower limb overuse injury are recognised as multi-factorial and diverse, it would appear logical that this must be reflected in formulating a screening tool for such injuries (88, 175). Such screening tools or protocols which acknowledge the multi-factorial nature of injury risk will be referred to here as functional screening protocols.

Development of such functional screening protocols, based upon the best available evidence, is considered important, not only from a preventative perspective but also from a vocational and clinical perspective. The actual cost of 'preventable' injuries to the person, their employer, and wider society can be difficult to determine. However, data obtained from ACC for the NZ army for the period 2004 to 2009 demonstrates significant potential savings. Preventable work related injuries in the lower limb (data for the same group of injuries as identified by Davidson et al (35)) had a prevalence rate of 10%. Sick pay cost estimates (with an assumption that most who suffered were on a lower end pay scale: non-commissioned officers with base level salary) amounted to approximately $9million NZ per year during this period. It is therefore considered important that potential injury screening protocols continue to be investigated and developed. The aim of this study is to address this issue, through investigation of a novel, functionally based screening protocol: FootFAST.

The overall aim of FootFAST is to provide a quantitative description of the movements occurring during the ground contact phases of bipedal motion at the foot. These are considered to be descriptive of the biomechanical forces acting on the lower limb, and thus potentially be indicative of overuse injury risk. Traditionally, a screening assessment is made

based on anatomical differences held in common by a group of injured persons; this is a deductive process, which has not been successful. As referred to above, various functional anomalies are seen in those with injury, but all persons with these differences do not subsequently have an injury (177). This problem has informed the development of FootFAST, which recognises that it is not the presence or absence of one functional or anatomical anomaly that is observed in injured persons, but several. Based upon three separate tests included within the FootFAST assessment, it is proposed that this screening tool has the potential for more accurate assessment of injury risk, as more than one positive test outcome is required to produce a result indicating a 'high' level of risk (67).

A previous study conducted by the author suggested that FootFAST may have predictive validity in the identification of overuse injury risk (1). Despite this encouraging result, the inter and intra-rater reliability, and criterion based validity of the three FootFAST tests has not been assessed, and this is an important pre-requisite for a clinical recommendation. Thus, the aim of the studies in the subsequent Chapters is to determine the inter and intra-rater reliability of the three tests within FootFAST, and also to assess the criterion based validity of each of the three tests. In order to assess criterion based validity, each test will be assessed against the best available comparison tests, i.e. software or biomechanical equipment-based assessments.

13.3 The FootFAST protocol

13.3.1 Test 1 (Hindfoot angle)

The subject stands relaxed in bare feet and the alignment of the heel is assessed in relation to its alignment with the horizontal. The scoring system is:

0. Ideal hindfoot alignment. The heel is perpendicular to the floor or less than 4 degrees eversion or inversion.
1. Slightly Everted. The hindfoot eversion is 4 to 8 degrees.

2. Excessive Eversion. The hindfoot eversion is greater than 8 degrees.

3. Inverted hindfoot greater than 4 degrees.

13.3.2 Test 2 (Single leg stance)

The subject stands on one leg with arms in their preferred position, attains balance with the eyes open and then closes the eyes. The examiner observes the test. The scoring system is:

0. Stable. The foot remains balanced, without perturbation or oscillations.

1. Mild Instability. On occasion, the foot tends to oscillate or correct itself.

2. Unstable. The foot is continuously oscillating and there may be a reduction in the ground-surface contact at times.

3. Laterally unstable. The foot continuously falls towards the lateral edge of the foot, losing ground contact on the medial aspect.

4. Cannot perform task: As indicated, if the participant is not balanced for at least five seconds.

13.3.3 Test 3 (Heel raise)

The subject stands on one leg and rises up onto the toes of the planted foot, with the eyes open. The scoring system is:

0. Smooth transition. The subject rises easily onto the toes, is stable and tends to roll forward.

1. Unstable. The subject can rise onto the toes but is unstable.

2. Resisted. The subject has difficulty rising onto the toes and may even tend to fall backward.

3. Inversion. The ankle tends to invert as the subject rises onto the toes. There is a significant risk of ankle sprains.

Subjects are assessed on each foot, and the worst score for either foot in each of the three tests is added up. This scoring system assigns the subject to one of three overuse injury risk groups.

0. Low Risk, a combined score of 0-2. Interventions are not required.

1. Medium Risk, a combined score of 3 to 5. Interventions recommended if the participant is in a high risk environment.

2. High Risk, a combined score of 6 or more. Interventions are required to avoid the occurrence of overuse injury.

13.4 A justification of the three tests comprising FootFAST (1)

There are three tests within the protocol: the biomechanical rationale for including each of these in a predictive model are as follows:

13.4.1 Test 1 (Hindfoot angle)

The test of hindfoot angle in upright, relaxed, weight-bearing stance was chosen as it was considered indicative of the nature of heel strike (Dr Charles Baycroft, Personal Communication), and this posture has been identified as predictive of the mean path of the hindfoot during the first sixty per cent of the walking cycle (159). The relationship between hindfoot angle and injury has been investigated previously with successful results: subtalar tilt of bony structures was shown to be significantly different between non injured (4.6 $\pm$ 3.1 degrees) and chronically injured persons (12.4 $\pm$ 9.6 degrees) (127), with no differences between males and females (178). This finding suggests a potential link between talar tilt anomalies either resulting in, or as a result of injury. Similar findings were observed by Beynnone et al (49), who concluded that tibia vara, talar tilt and calcaneal inversion (179) were separately and together seen in higher risk injury groups.

Heel strike during gait has been previously identified as being related to overuse injury (72, 88, 119), where an ‘optimal’ contact position would be one that promotes a rapid leg alignment and less compensatory movements. Studies have shown that increased loading to either the extreme medial or lateral border of the foot is apparent in those with overuse injury (88). Stolwijk (72) found that those whose weight tended laterally at the foot tended to experience patello-femoral pain (no relative risks reported); those who tended to load more medially experienced significantly more exercise related lower leg pain (P=0.007); while those with a pronounced lateral roll off on to the forefoot had significantly more running injuries (P=0.031) (119).

Studies which have investigated function found similar results: increased calcaneal eversion was identified as one of two risk factors for ankle ligament injury (increased ankle ROM was the second) (49). A preliminary study performed by Willems et al. (119) which attempted to identify potentially injurious biomechanics during gait also observed that eversion during ground contact was a risk factor, as it results in increased lateral roll-off (specifically, eversion during forefoot flat p=0.003 and heel off p=0.049). These studies are potentially limited in some respects in terms of their generalizability, in that they both used college athletes or physical education students as participants, and that they included a large amount of variables in the statistical analysis (thus decreasing the power). However, the consistency and nature of their findings suggest that the results may be somewhat generalizable.

There is a lack of consensus in the current literature as to what constitutes a ‘normal’ inverted or everted rearfoot pattern. While there is consensus in terms of the overall concepts, there is much controversy over what defines the parameters of each. It is not within the scope of this study to define what should be categorised as such, however, for the purposes of the current study, the parameters reported by Willems et al (180) will be used: a deviation greater

than four degrees from vertical being classified as either inverted or everted respectively. This is for two main reasons: firstly, their definition coincides with a biomechanical framework for classification of injury risk, and secondly, their study utilises the same foot scanning equipment and software as will be used in the current study.

Based on the discussion above, on a structural, functional, and biomechanical level there is evidence that the test of hindfoot angle could have merit in the prediction of overuse injury. Hence, the choice of including the test of hindfoot angle as part of the FootFAST protocol has been justified.

13.4.2 Test 2 (Single leg balance)

Standing leg balance, also known as the "single leg stance test" is one of many standing tests used by practitioners. It is described by some as the most common clinical assessment tool (160) (and is commonly used by researchers) to assess both postural and functional stability. This functional stability evaluation is incorporated as the second test within FootFAST. There are several theories which underpin the investigation of stability as the basis of an overuse injury screening test. These are based on the association between stability and proprioception; when referring to stability or instability, it is essentially as a measure of proprioception; dynamic instability and proprioceptive deficit are said to go hand in hand (181).

Proprioception is defined by Bonnel (182) as the system that comprises receptors, pathways, and nerve centres in both conscious and unconscious perception of the relative position of the body parts in relation to one another. The research literature on proprioception and injury rates has been fairly consistent (refer to Chapter 2): those who score low on various proprioceptive tests (active and passive core and trunk proprioception) have been shown to be at increased risk of injury to the lower limb, and in particular the knee (100, 154). In addition, those who complete proprioceptive training programmes have been shown

to reduce their occurrence of future injuries by around 35%, based on evidence gained from RCT's and intervention studies (refer to Chapter 2) (35, 104).

Interestingly, it has also been found that those who did not sustain an injury (i.e. were at low risk of injury to start with) did not benefit from proprioceptive training (no increase in test scores) (104). This adds further weight to the argument that proprioceptive disturbance is relevant to injury occurrence, with a link between deficits in proprioception and overuse injury, although cause and effect is yet to be established. This can be explained through the manifestations of a proprioceptive deficit: by definition, low levels of proprioceptive capacity lead to decreased motor control, due to physiological deficiencies at both the central and peripheral levels (183). Interestingly, these same physiological deficiencies are observed when a person is in a fatigued state (184); therefore, it could be suggested that those with low levels of proprioception are in a sense 'pre-fatigued' with respect to co-ordination. Fatigue was discussed as an injury risk factor in Chapter 2.

Anecdotal clinical evidence suggests that the most frequently cited reason for chronic injuries of the lower leg is 'postural instability', defined as the feeling or demonstration of unsteadiness during standing or bipedal movement (185). Such an argument suggests that screening for lower leg injury requires assessment of stability; indeed in more recent literature, this has been achieved through an assessment of proprioception (186). Despite the common assumption that instability and proprioception are considered complementary, this may not provide a complete picture. Proprioception performance does not in isolation indicate full 'stability'; both anatomical and functional variables have an effect on overall stability (182). Anatomical factors may include ligament laxity and bony congruence, while functional factors refer to the quality and onset of active muscular contractions. Proprioception testing typically involves the use of joint angle replication and passive position sensing, which arguably only truly assess functional aspects of stability (182).

Because of this, the current screening test has been designed to assess both the functional and anatomical aspects of stability with a single leg balance test.

Single leg balance tests are recognised in the literature as a valid measure of injury susceptibility (160, 182). A decreased ability to maintain single or bilateral stance has been shown to be associated with increased injury risk for the individual (56, 59), as maintaining single leg stance requires control of weight support, vertical alignment of the body segments, and total body equilibrium (187). Beynnone et al (49) found that anterior-posterior sway when standing on one foot with eyes closed was predictive of ankle injury occurrence. This test is very similar to that used in the current study. Beyond this, Schneiders (160) and Bonnel (182) both similarly concluded from their studies that sex was not a confounding variable in performance of this test, and the same is true for limb dominance (188). These findings provide further support to the use of a standing single leg balance test in the current protocol.

Broadly, there are two means of conducting a balance test. There are those which require expensive equipment and software, such as the use of force plates for balance assessments, where the difficulty of keeping the centre of mass above the centre of pressure is reflected in the variability of the ground reaction forces, stabilometry, and centre of pressure excursions (56, 187, 189). Then there are those which are based upon purely visual observations. These are by far the most frequently used, particularly in clinical settings such as neurology, paediatrics, geriatrics, orthopaedics, and sports medicine (190). However, very little is currently known regarding the reliability and validity of such visual tests (58) and thus visually judged tests are the focus of the current investigation.

When designing the single leg balance test for the current study, many currently used assessments were considered (Dr Charles Baycroft, Personal Communication). Various visually-assessed standing balance tests are currently used with high success with respect to

their psychometric properties (56, 153, 191), the Rhomberg test being recognized as the traditional means of identifying functional instability. Since the original development of the Rhomberg, testing processes have expanded, and a range of tests are now available for the assessment of functional instability for specific purposes, including:

1. Standing balance assessments of both dynamic and static abilities, where the initial part of balancing is seen as a dynamic phase, which leads unto a static phase (191).
2. The Brunel Balance Test, which has been reported to have scored 100% for both intra and inter-observer reliability (191). (This test is mainly for assessments post-stroke and as a fall prediction tool for the elderly, but it is helpful as it illustrates that standing balance tests are a reliable and feasible option.)
3. The balance error scoring system (BESS) assessment, another observational method of evaluating stability, with the addition of a predefined set of movement errors which also contribute the final score (189).

As has been made apparent, there were many test protocols available. When developing the current assessment, the need to maintain high levels of reliability and validity was considered imperative. Reliable tools should be sensitive enough to detect any differences or anomalies which could not have been due to chance. This is directly related to the standard error of measurements, and therefore to the sensitivity of the assessment (57). It has been found that when assessing single leg stance in terms of time (duration) only, the repeated test reliability is poor; Goldberg et al (57) found that their repeated testing of single leg balance time had a measurement error of 4.8%, (or 24.1 seconds), therefore the researchers concluded that the reliability of the test was poor. The results of that study suggest that both measurements of 'balance time' and/or repeated performance measurements are not feasible in a reliable test.

Even in cases where assessments are made against standardised or expected values for balance time, (i.e. where the absolute time is compared to a standardised result), a range of

problems have been identified. For example, various papers have investigated the validity of current normalized single legged stance times (eyes open and/or eyes closed) (190). The best *average* time for a substantial cohort (n=98, 18-39 year olds) was found to be 15.2seconds, which is an interesting finding given that the most common assessment time is twenty (189) to thirty seconds (190)**.** In fact, single leg stance times are highly variable (even for an individual as first trials are considerably shorter than subsequent trials) and consequently reliability is poor; as a result, this test is not recommended for concussion assessments, repeated testing, or for monitoring an intervention (189). The above points have been considered in the FootFAST protocol, and single leg stance is primarily measured in terms of quality of performance, not quantity of time.

Hence, it was then proposed that an investigation of steadiness during single leg stance would provide more reliable and valid results. In order to assess the criterion based validity of the visual measurement, a pressure plate recording of the single leg stance test would be used for comparison. Although a number of investigations of steadiness have been undertaken in the past (187, 192, 193), low scores of criterion based validity (low agreement with criterion measures) have previously been found (58). Normally, the velocity and positioning of the centre of pressure under the foot is used (assessed quantitatively using specialist equipment), and compared with the time performance of the individual. A comparison of the centre of pressure velocity and movement with a visual assessment of performance quality (as will be utilised here) has not previously been reported in the research literature. This is another and alternative means through which the results of this study are novel, and if the validity seen here is improved, it may inform a change in practice.

In order to further address the poor reliability found in past studies, recommendations from previous researchers have been taken into consideration in the design of FootFAST. It has been suggested that the low reliability estimates found in single leg stance tests (i.e. wide

performance variability) are because the task fails to challenge the participant sufficiently (59). In order to address this, the FootFAST protocol requires the participant to keep their eyes closed; this has been proven to be significantly harder (based on test scores) than when assessed with eyes open (188) which has been observed to produce a floor effect. Secondly, Beynnone et al (49) investigated alternative possible scoring methods for the single limb balance assessment in order to identify the optimal methodology for both inter and intra-rater observational scores, including comparisons of performance classifications generated from force platform outputs. They found that a high reliability was attributed to the use of a highly discriminating rating scale (eleven points versus four points as previously used clinically). In their study, raters were asked to view minimal aspects of performance at any one time, clear operational definitions were given, and there was a short pre-training session for the raters which outlined levels of performance using a video recording (49). Aspects from Beynnone et al's (49) protocol, which was proven to be effective in creating an environment conducive to acceptable reliability scores, has been directly used in the design of the FootFAST protocol.

13.4.3 Test 3 (Heel raise)

This test is designed to be an indicative measure of a critical part of the bipedal cycle: the push off. The main goal of bipedal movement is forwards momentum, which is almost completely achieved by the action of plantarflexion: applying a force great enough to push the body weight forwards. In addition to this, a smooth and efficient transition into plantarflexion from stance is logically beneficial from a biomechanical point of view, in order to maintain forwards motion and energy efficiency (refer to Chapter 2).

In order to achieve 'good' plantar flexion, the subtalar joint has to supinate to allow the toes to dorsiflex as the heel lifts. Thus, the ability to plantarflex is related to supination resistance and the activation of the windlass mechanism (194), linking the hindfoot and

forefoot. People who have difficulty rising onto the toes of one foot are very likely to have high supination resistance and impaired windlass activation (and possibly an oblique subtalar joint axis (195)). This is important in the assessment of injury risk, as the association between impaired windlass activation and various foot injuries and pathologies is well documented (84, 194).

Based on the information outlined, it would seem that the toe off phase is arguably the potentially most injurious phase of weight bearing during bipedalism. This may be due to the large amount of force which must be applied through the foot while it is in this position, in order to provide sufficient propulsion to move the body forwards.

A stiffness in the ankle joint (found to be a reliable and valid measure of plantarflexion weakness (196)), a tightness in the gastrocnemius muscle (the main plantarflexor) (197), or generalised weakness in the plantarflexors (198) can compromise plantarflexion . As indicated above, such a situation will compromise an efficient movement pattern, and therefore will be counterproductive in achieving efficient gait. The main effector muscle of plantarflexion, which is activated during toe-off, is the medial gastrocnemius. Raising onto the toes on one leg can therefore be considered a very close approximation to the same movement, as medial gastrocnemius (along with the soleus) is the main effector muscle for both (199-201). This provides the justification for a type of 'forefoot propulsion test' being chosen as the third test within the protocol.

13.5 Final comments

Finally, it has been consistently shown that repeated performance of trials such as single leg stance elicits a practice effect, and this effect is even more pronounced in novel tasks, such as in this case, performing a heel raise (189, 202). For this reason, only one performance and no practise attempts will be employed as the standard procedure for all three tests included in this study (as detailed in the methodology presented in Chapters 6 and 7).

Despite the justification of the three tests described above, the reliability and validity of such visual tests has never been investigated. The aim of the current study is to assess the inter and intra-rater reliability, and criterion based validity of the three tests which comprise the FootFAST protocol. The results of these investigations are reported in the feasibility study (Chapter 6) and then the main study (Chapter 7).

13.6 In the next Chapter

The apparent complexity of the FootFAST protocol dictated the necessity of a pilot trial to assess the data collection method, and also the resources involved in terms of personnel, time, and space. After some preliminary pilot testing, a feasibility study was devised. Feasibility studies, as defined by Arain et al (203), have the primary purpose of measuring the critical parameters needed for power calculations to design main research investigations- determining the dispersion of the results and calculating a suitable sample size. In addition, a feasibility study answers additional questions regarding utility: for example whether the timeframes, costs, quality of data and quality of outcome justify the methods selected (204-206). For the current research, this involved determining the suitable size of confidence intervals for agreement results, in order to estimate sample size requirements.

The next Chapter presents the feasibility study, which incorporated the same methodological procedure as was to be used within the main study (Chapter 7). The feasibility results are analysed and reported upon, with the aim of identifying potential problem areas to be addressed prior to the main study. This helped to ensure that the main data collection was of a high standard methodologically (i.e. high internal validity).

14 Chapter 6. What are the psychometric properties of the injury risk assessment tool: FootFAST? A feasibility study

14.1 Overview of Chapter

The first part of the thesis comprised narrative reviews (Chapters 2 and 3) and a systematic review (Chapter 4) of current literature within the field of injury prevention screening, and interventions for lower limb overuse injuries. Findings from these reviews were utilised in the design of the methodology utilised in the current feasibility study. This Chapter presents a feasibility study which was used to inform the design and power analysis (determining participant numbers required) for the main study. It is the first step required in order to answer the final three research questions (refer to Chapter 1, p.14) for this thesis, which were:

3. What is the inter-rater reliability of the three tests within the FootFAST screening protocol?
4. What is the intra-rater reliability of the three tests within the FootFAST screening protocol?
5. What is the criterion based validity of the three tests within the FootFAST screening protocol?

The Chapter begins with an outline of the methodology, then the results are presented and findings are discussed.

14.2 Introduction

The feasibility study was conducted using the intended methodology planned for the main study, with the additional purposes of answering the following research questions.

1. Is an intra-rater reliability assessment of the three FootFAST tests feasible?
2. Is an inter-rater reliability assessment of the three FootFAST tests feasible?
3. Is a criterion based validity assessment of the three FootFAST tests feasible?
4. What is the size of study required in order to produce confidence intervals of K<0.2 with 80% power?

14.3 Methods

This is a repeated measures cohort with a participant and assessor sample of convenience. Ethical approval was obtained from the University of Otago Research Ethics committee and Maori consultation was also completed (refer to Appendix 9).

14.3.1 Participants

Participants consisted of ten New Zealand Army soldiers, nine males and one female, aged between nineteen and twenty five years. None of the participants had an injury that was current, or had experienced an injury during the previous three months that affected their normal ambulation. All participants were volunteers from a local reserve unit, and were approached through the unit's officers; all provided informed consent and signed the consent form (refer to Appendix 2).

14.3.2 Inclusion criteria

Participants were selected from a group of New Zealand army territorial force soldiers who were present at training on the night of the 19th September 2012. The first ten soldiers who volunteered in response to the invitation to participate, and who met the inclusion criteria, were selected.

14.3.3 Exclusion criteria

Participants were excluded if they had a current injury, or one in the previous three months that affected their normal ambulation.

14.3.4 Assessors

Assessors consisted of six clinically trained (Bachelor of Physiotherapy) and clinically experienced (minimum of three years of clinical practise) physiotherapists from the School of Physiotherapy, University of Otago; all were currently practising and registered by the Physiotherapy Board of New Zealand. A recruiting email was distributed to staff across the School, and the first six voluntary respondents were chosen for the study. All assessors signed the consent form (refer to Appendix 3).

14.3.5 Inclusion criteria

To be included as an assessor in this study, the assessor had to be working for the School of Physiotherapy Clinics – University of Otago, during September 2012 (in order to receive the recruitment email, refer to Appendix 1). Assessors had to have had a minimum of three years clinical experience as it was identified in the literature review (Chapter 3) that clinicians with this level of experience were optimal for producing consistent results (207). The first six respondents who met the criteria were selected.

14.3.6 Exclusion criteria

Assessors were excluded if they had previous experience using FootFAST, in order to standardise the amount of experience in using the protocol under investigation.

14.3.7 Equipment

Two high speed digital video cameras (Sony Alpha A37 DSLR / SLT; Sony Limited, New Zealand) were used to capture still images of the hindfoot, and videos of both single leg

stance and heel raise performances. Such images and videos were used by assessors for later analysis. The RS Scan™ mat (RS Scan International, Belgium) and accompanying Gait™ software were used to capture the centre of pressure data and underfoot pressure characteristics of the foot during single limb stance and heel raise performances. The Silicon Coach™ (SiliconCOACH, New Zealand) software programme was used to calculate hindfoot angle. A black permanent marker was used to outline anatomical reference points described below. One 1.5 metre measuring tape was used to calibrate the camera and measure the distance between the ischial tuberosities. Finally, nine reflective markers, 10mm in diameter, were used to mark reference points for anatomical angular measurement. Specifically, these were placed bilaterally on each malleolus, the medial condyle of the tibia, the lateral aspect of the head of the fibula, also on the palpable tip of the calcaneus, and a marker was placed on the Achilles tendon at the fusion of the gastrocnemius muscles with the triceps surae.

14.3.8 Kinematic data

For the criterion based validity comparison of Test 2 (single leg stance) and 3 (heel raise), assessor results were compared with centre of pressure data. Both assessments (visual and objective), were completed during the same performance. The centre of pressure movements and sway path were tracked using the RS Scan™ mat (RS Scan International, Belgium) and accompanying Gait™ software. The data were viewed within the Gait™ software package, and the centre of pressure path translated onto an image of the plantar surface of the foot in contact with the mat.

14.3.9 Procedure

In order to complete the data collection phase, each participant completed the following procedure in turn. The environment was set up as in Figure 14 below.

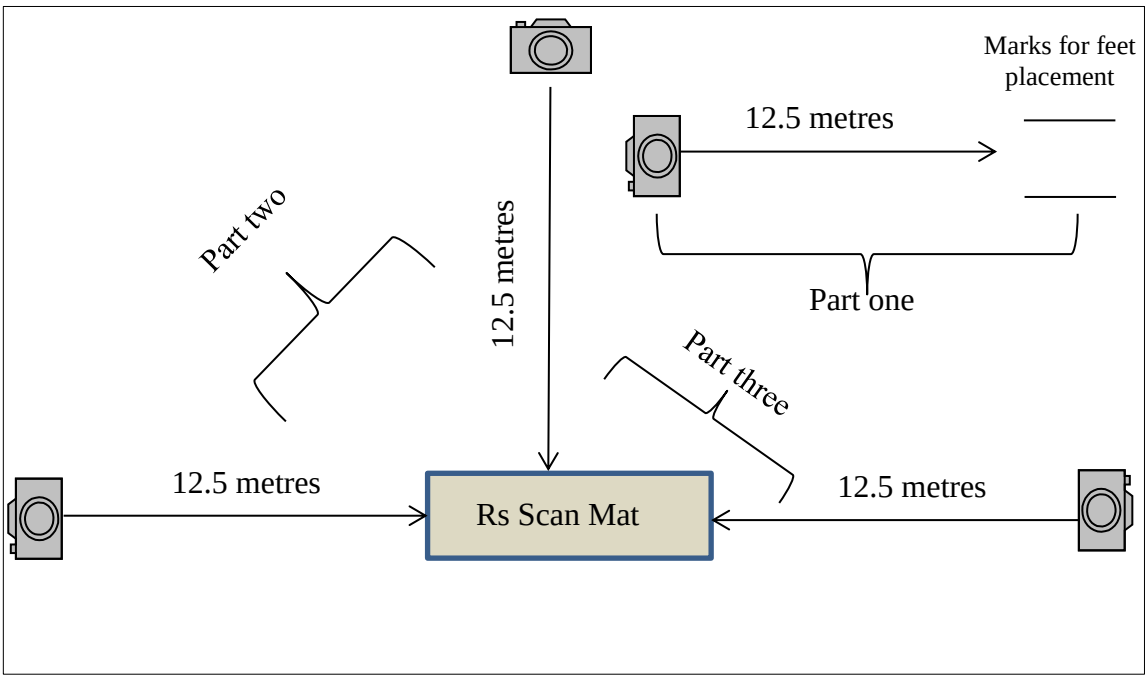


Figure 14: Configuration of the data collection environment.

There were three parts to the procedure, corresponding with the three tests within FootFAST. For the first Test, the environment was set up with a camera placed as far from the viewing field as possible to minimise for parallax error, and zoomed in so that minimal excess picture was included (208). For the second and third Tests, the environment was set up as follows: the RS Scan™ was placed in the centre of the room so that the participant could not fall onto any other piece of equipment, or use a wall for support. Two cameras were set up to record images either from the medial side and from the front (Test 2) or from the lateral side and from the rear (Test 3) of the planted foot. These cameras were situated as far from the mat as possible (12.5 meters away), and perpendicular to the mat to avoid parallax error. A right angled triangle was created in order to ensure that the camera was perpendicular to the mat, with sides measuring 3m / 4m / 5m (according to Pythagoras' Theorem). The foot and the lower half of the shank were included in the field of view (as shown in Figure 16 below).

14.3.10Test 1 (Hindfoot angle)

The participant was in bare feet and wore clothing above knee level for this trial. Two separate digital images were taken, one with joint markers attached- for reference points when measuring the angles within Silicon Coach^{TM} (SiliconCOACH, New Zealand), and one without for visual assessments. Silicon Coach^{TM} (SiliconCOACH, New Zealand) has beenA distance reference line 10cm in length was drawn on the wall, within the field of camera view and perpendicular to the floor. This distance was used for scalar calibration of these images using CAD software program Silicon Coach^{TM} (this is in accordance with the calibration process outlined in the software's reference manual - SiliconCOACH, New Zealand).

The distance between the ischial tuberosities (identified through palpation) was measured using the 1.5 metre measuring tape; this was then apportioned as the distance that the feet were placed apart. Lines were drawn on the ground, perpendicular to the wall, to indicate the distance apart as measured between the ischial tuberosities. The participant placed both feet on these lines, with the line running through the middle of their heel and under the second toe. Although this may not be indicative of typical stance, it minimises parallax error as the heel is optimally aligned with the optical field of view. (209, 210)

The lower limb from above the knee joint distally was included in the field of view and a digital image was captured (as shown in Figure 15). Once an initial image was captured, joint markers were placed on the participant. These markers were placed in similar locations as used by Davison et al (35), as outlined above. Marks which were 2cm in length were drawn with a permanent marker on both lateral and medial borders of the calcaneum where it made contact with the ground. The participant was returned to the same position as for the first image, and a second image was captured.

Figure 15: Example of the images captured during Test 1.

14.3.11 Test 2 (Single leg stance)

All marks and markers were removed from the participant. The participant then stood on the RS Scan™ mat with the axis of the foot parallel with the longest side of the mat, and was permitted to position their upper body and arms in their preferred position. When instructed to, the participant lifted one foot off the ground, closed their eyes and attempted to remain balanced for as long as possible. Once the participant could no longer maintain balance the trial was accepted as being complete. There were no practice attempts permitted and no repeated trials, as this is the procedure used within FootFAST (1). The ankle joint was the focal point of the footage obtained in this performance, as it has been shown to be the joint of primary importance during single leg stance (211).

Figure 16: A still frame from a single leg stance video captured during Test 2.

14.3.12 Test 3 (Heel raise)

Participants again stood on the RS Scan™ mat and were permitted to hold their upper body and arms in their preferred position. When instructed to, participants lifted one foot off the ground and attempted to rise unto the toes of the planted foot. The participant was asked to maintain this posture for about three seconds. There was only one attempt allowed for this test and no repeated trials, as this is the procedure used within FootFAST.

14.4 Data collection and analysis

The images captured for Test 1 (the hindfoot position of each participant) without bodily markers along with the video footage captured for Tests 2 and 3 were edited to ensure minimal unnecessary footage or background images were present. These were then included in a DVD, and sent to the six assessors. The videos were in identical sequential order for each assessor, and the same order sequence was used for the repeated assessments.

The assessors used this visual information to grade the participants using the FootFAST procedure. All written information provided to the assessors is included in Appendix 4. The assessors were all asked to consider any recommendations they might have for the main study with regard to the information with which they were provided. (The issues raised are

discussed in the 'Discussion- Further implications for the main study' section later in the Chapter.)

14.4.1 Continuous and categorical scales

Results obtained from the assessors were analysed as both continuous and categorical data sets. For the continuous data set, assessors were required to place a mark anywhere on a visual analogue scale which corresponded to their assessment of the participant. The exact position of these marks were analysed as continuous data. As FootFAST is designed to elicit a categorical response, these marks were also transformed into a categorical data set. The category closest to the mark placed on the visual analogue scale was identified as the category assigned by that assessor. For example, if the mark was placed between 'stable' and 'mildly unstable' in Test 2, but closer to 'stable', then the outcome of the assessment was 'stable'.

The justification for analysing the results as both categorical and continuous data was as follows: A recent study (2012) by Brosseau-Laird et al (212) investigated the application of visual analogue scales to categorical data, as performed in the current study, and the effect this had on the data and the corresponding analysis. They found that, in general, categorical data is more sensitive to sub-optimal sample sizes and violations of the assumption of data normality. Conversely, continuous data was more sensitive to asymmetrical category thresholds and was highly susceptible to bias in cases of large factor loadings (212). On a whole, it was concluded that in cases where there are fewer than five categories, categorical data sets should be used, and for five or more, continuous data sets are preferable.

Taking this previous work into consideration, it was still unknown whether a continuous or categorical data would be optimal. The use of a categorical scale was not ideal as the data was not known to be normally distributed, and for some of the Tests there were five or more categories of response (Test 2). However, the use of a continuous scale was also

flawed due to the presence of asymmetrical category thresholds, and for some Tests there were fewer than five categories of response (Tests 1 and 3). Given this dilemma, it was decided that the data would be interpreted as both categorical and continuous sets, and a decision would be made based on the findings.

14.4.2 Inter-rater reliability

Each assessor (N=6) assessed each participant (N=10) in each of the three tests within FootFAST. The inter-rater agreement was calculated for each assessor with each participant, and then the agreement levels were averaged. An example of this calculation provided below.

Using a hypothetical example of Participant 1 -

Each of the six assessors gave a score to Participant 1 for all of the three FootFAST tests. The inter-rater agreement between all assessors for Tests 1-3, for participant 1 were as follows – K=0.22 (Test 1), K=0.34 (Test 2), K=0.43 (Test 3) respectively. This calculation was repeated for all ten participants, after which, results were averaged for each of the three tests separately.

14.4.3 Intra-rater reliability

Three assessors (N=3) assessed each participant (N=10) in each of the three tests within FootFAST three times. Thus, each participant individually was allocated nine scores from each assessor. An example of the intra-rater agreement calculation is given below.

Using a hypothetical example of Assessor 1, Participant 1:

The three scores allocated on each of the three repeated trials were:

Test 1 - 1, 1, 1;

Test 2 - 2, 1, 2; and,

Test 3 - 3, 1, 2.

The intra-rater agreement may then be, for this example:

Test 1: K=1.0;

Test 2: K=0.64; and,

Test 3: K=0.32.

This process was repeated for the three assessors who completed repeat assessments and all ten participants. Intra-rater agreement scores were then averaged across the three assessors for each test separately.

14.4.4 Criterion based validity

For Test 1 (Hindfoot angle), the images taken with body markers were assessed using the Silicon Coach™ program. Silicon Coach™ has been identified as a reliable and objective measure of joint angles (19), with intra and inter-rater reliability of ICC=0.92 (213). Further, its use with 2D video analysis has excellent correlation with 3D motion capture measures (213). The anatomical reference markers, along with a vertical reference line for image calibration, were used to provide reference values for the horizontal calculation of hindfoot angle. These angles were then compared with each of the angular results from the assessors (an average of the three responses). An average agreement level (Kappa) was then calculated across all ten assessors.

The data captured by the RS Scan™ device during Tests 2 and 3 were assessed within the Gait™ software package. For Test 2 (single leg stance), the centre of pressure movement ('COP', millimetres) was tracked within the software system. The most extreme excursions of the centre of pressure were identified, also whether the excursion was to the medial or lateral aspect. This was chosen as it is presumed to be that which best represents what the assessors were basing their grade on: the visual assessment of stability is essentially the monitoring of how the body makes constant adjustments in order to maintain the centre of gravity above the COP (refer to Figure 17). Better balance, which is viewed as being less excursions, is indicative of better proprioception (155). The agreement between results

gained from the Gait™ software and the averaged result (as these were assessed on three occasions) from the assessors for each participant separately were calculated, and then an average agreement level was calculated.

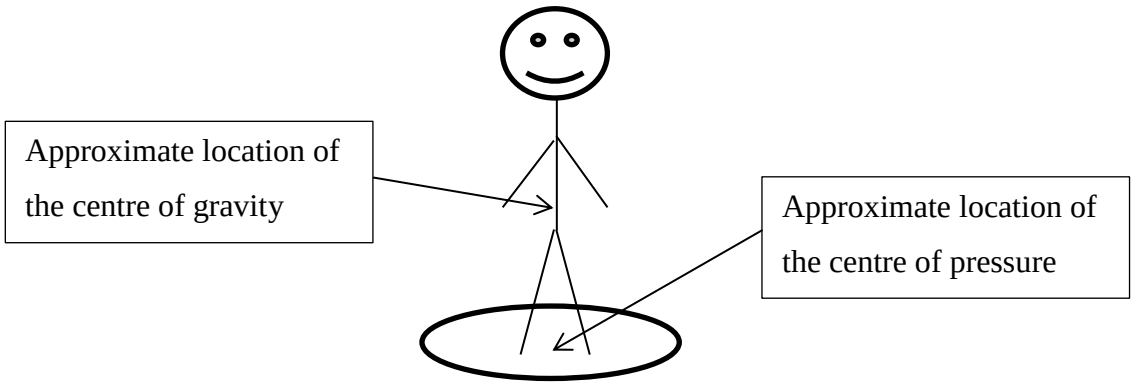


Figure 17: Illustration of the concepts of centre of gravity and centre of pressure.

Figure 17 illustrates the approximate locations of the COP and the centre of gravity on a person, while standing in the position indicated. The centre of gravity is essentially the body's pivot point, and is located at the exact point were equal weight (and thus momentum) is acting in each direction. The COP is the point, where equal downwards pressure is acting on all sides (approximately centralised either between the two feet when bipedal, or in the centre of the plantar surface when unipedal). Balance can be simplistically referred to as the maintenance of the centre of gravity directly above the COP. When a person becomes off-balance, even by almost negligible amounts, the centre of pressure will move in the same direction of the sway. The speed of the COP movement is also proportional to the speed of the centre of gravity. Thus, monitoring the COP is indicative of the movement of the centre of gravity. (214)

14.4.5 Feedback from assessors

As the primary aim of the current feasibility trial is essentially to 'test' the methodology planned for the main study, it was considered appropriate to include a verbal debrief with the assessors. This consisted of either a face-to-face meeting or phone call (depending on assessor proximity), and assessors were asked the question:

"Do you have any feedback which may improve the quality of the main study, or improve the ease of information passage between the instruction sheets provided and an assessor?"

All responses were noted and presented in the Discussion section below.

14.5 Statistical Calculations

All statistical analyses were conducted using STATA 7.0 software (215). Dr Andrew Gray (Biostatistician, School of Preventive and Social Medicine, Dunedin School of Medicine) was consulted with regards to choice of appropriate statistical test, and confirmation of results.

14.5.1 Inter-rater reliability

For inter-rater reliability calculations, Fleiss's weighted kappa (FWK) test was used. This is represented on a scale from -1.0 (perfect disagreement) to 1.0 (perfect agreement) (216).

Results gained from both continuous and categorical data sets were used to make comparisons between assessors. The inter-rater agreement was calculated for each participant individually, and then averaged. Results will be presented for each of the three tests separately.

In addition, the level of inter-rater agreement for each participant was also investigated. This sought to answer the question whether the individual participant (or in ecological settings, the patient), had an effect on the ability of assessors to produce consistent results.

14.5.2 Intra-rater reliability

To calculate intra-rater reliability, the results given on the three separate occasions of testing were compared using Cohen's weighted Kappa. Specifically, the agreement level for each of the three assessments (Tests 1, 2, and 3) of each patient individually were calculated, then averaged across all participants, for all assessors who completed repeat assessments (N=3).

14.5.3 Criterion based validity

To calculate the criterion based validity between measurements made by assessors and the objective comparison, the results produced were averaged for each participant individually for each of the three tests. These averaged results for each test were then compared to the results from the RsScan Gait^{TM} software analysis (Tests 2 and 3), and from the Silicon Coach^{TM} (Test 1). Cohen's weighted Kappa was used for this statistical comparison.

Kappa values obtained were interpreted based on the limits set by Sim and Wright (216) for Kappa <0 poor, 0.01-0.20 slight, 0.21-0.40 fair, 0.41-0.60 moderate, 0.61-0.80 substantial, and 0.81 -1.0 almost perfect. Clinically acceptable levels of Kappa are taken as K > 0.4 (216, 217).

14.6 Results

It is noted that this study was not powered, and thus the results presented cannot be used to derive definite conclusions from.

Results for inter-rater reliability are presented for both categorical and continuous variables in Table 14-3 and Table 14-4. All comparisons produced a score which corresponds to 'slight' agreement (216). Intra-rater reliability scores of all three tests for all assessors were better: all comparisons produced an agreement level which was above the clinically acceptable minimum of Kappa>0.4 (Table 14-5). Criterion based validity comparisons

produced mixed results: Kappa values of 0.700 and 0.804 were found for Tests 1 and 2 respectively; whereas Test 3 resulted in an agreement of K=0.028 (
Table 14-6 14-6).

Table 14-1 shows the characteristics of the ten participants (nine males and one female), chosen as a convenience sample from a New Zealand army training night. Based on other studies utilising the same population, these participants seem to be similar in terms of the demographics reported (80).

Table 14-1: Participant characteristics (N=10).

Characteristic	Mean (SD)
Height	177cm (6.3)
Weight	82kg (10.9)
Age	19yrs 9months (13 months)
Ethnicity	Maori (N=5)
	Pacific Islander (N=1)
	Pakeha/NZ European (N=4)
	Asian or other (N=0)

SD: standard deviation

Table 14-2 shows the characteristics of the assessors who were included in the study (N=6). Three of the six assessors were musculoskeletal specialists, and four of the assessors were currently practising clinically, while two were lecturers and practising.

Table 14-2: Assessor characteristics (N=6) all with a minimum of three years clinical experience.

Assessor Number	Current specialist area	Currently practising
1	Musculoskeletal	Yes
2	Mackenzie methods	Yes (lecturing)
3	Musculoskeletal	Yes
4	Neurology	Yes
5	Musculoskeletal	Yes
6	Neurology	Yes (lecturing)

Inter-rater reliability

Table 14-3Table 14-3 shows that across all three FootFAST tests, the Kappa values for inter-rater agreement were only ‘slight’ for categorically assigned results. FootFAST is designed to be scored based on categorical outcomes (as outlined in Chapter 5). The confidence intervals for Test 2 include negative values (indicating agreement observed here was worse than would be expected by chance).

Table 14-3: Fliess’ weighted Kappa for inter-rater reliability of categorical data.

Test	1 (Hindfoot angle)	2 (Single Leg Stance)	3 (Heel Raise)
Fleiss’s weighted kappa (width of 95% confidence interval)	0.199 (0.044 to 0.353)	0.018 (-0.118 to 0.153)	0.164 (0.028 to 0.299)

Table 14-4 shows that the FWK calculated for the inter-rater reliability for each of the three tests within FootFAST ranged between poor and slight. The negative scores seen for both Tests 2 and 3 indicate that the inter-rater agreement was worse than would be expected by chance, as a Kappa of 0.0 indicates a level agreement corresponding to ‘chance’ (216).

Table 14-4: Fliess’ weighted Kappa for inter-rater reliability of continuous data.

Test	1 (Hindfoot angle)	2 (Single Leg Stance)	3 (Heel Raise)
Fleiss’s weighted kappa (width of 95% confidence interval)	0.088 (0.001 to 0.174)	-0.022 (-0.093 to 0.048)	-0.026 (-0.084 to 0.032)

Intra-rater reliability

Three of the six assessors completed repeat assessments.

Table 14-5 14-5 shows that Kappa scores for Assessor 1 ranged from ‘perfect’ reliability (Tests 1 and 3), to ‘fair’ reliability (Test 2). Assessor 2 had reliability scores which ranged from ‘moderate’ (Test 1) to ‘almost perfect’ (Test 3). Intra-rater reliability for Assessor 3 was consistently ‘substantial’ for all three tests within FootFAST. In general, Test 3 produced the most consistent results (two assessors scored ‘almost perfect’, one scored ‘substantial’). Test 2 produced the worst scores, from ‘fair’ (one assessor) to ‘substantial’ (two assessors). Finally, Test 1 scores ranged from ‘moderate’ (one assessor) to ‘almost perfect’ (one assessor).

Table 14-5: Cohen’s weighted Kappa (CWK) calculations for intra-rater reliability among the musculoskeletal physiotherapists, and average results, shown are for each test individually.

Assessor number / FootFAST test number	**CWK level of agreement (95% confidence interval width)**
1 / 1	1.000 (1.000 to 1.000)
1 / 2	0.348 (0.198 to 0.448)
1 / 3	1.000 (1.000 to 1.000)
3 / 1	0.565 (0.475 to 0.656)
3 / 2	0.667 (0.564 to 0.770)
3 / 3	1.000 (1.000 to 1.000)
5 / 1	0.667 (0.564 to 0.770)
5 / 2	0.677 (0.572 to 0.783)
5 / 3	0.643 (0.534 to 0.752)
FootFAST test	**Average CWK agreement**
1 (Hindfoot angle)	0.744 (0.680 to 0.809)
2 (Single Leg Stance)	0.564 (0.445 to 0.667)
3 (Heel Raise)	0.881 (0.845 to 0.917)

14.6.1 Criterion based validity

Table 14-6 14-6 presents results for the criterion based validity assessments. For Tests 1 and 2, there was 'substantial' to 'almost perfect' agreement with the objective comparison. However, for Test 3, the FootFAST test had only a 'slight' agreement with the objective comparison.

Table 14-6: Cohen's weighted kappa (CWK) values for the three tests within FootFAST, calculating criterion based validity (comparison of assessor results with Silicon Coach™ (hindfoot angle) and Gait™ (Single leg stance and Heel raise) software results).

Test	1 (Hindfoot angle)	2 (Single Leg Stance)	3 (Heel Raise)
CWK (width of 95% confidence interval)	0.804 (0.711 to 0.897)	0.700 (0.657 to 0.743)	0.028 (-0.063 to 0.119)

14.6.2 Musculoskeletal clinicians

Table 14-7 was constructed in hindsight following the observation of trends; it became apparent that reliability scores among the musculoskeletal physiotherapists were better than the results from the entire group. Results presented in Table 14-7 indicate that inter and intra-rater reliability of the three tests within FootFAST for the musculoskeletal physiotherapists ranged from 'fair' to 'almost perfect'. The width of the confidence intervals indicate that more data, specifically, a greater number of participants and/or assessors are required to produce a more certain result.

Table 14-7: Inter and intra-rater reliability, among only musculoskeletal clinicians, assessed using Fliess' linear weighted Kappa scores.

Comparison	Kappa	95% Confidence Interval
	Inter-rater reliability	
Test 1	0.444	0.200 to 0.6886
Test 2	0.373	0.085 to 0.660
Test 3	0.686	0.359 to 1.000
	Intra-rater reliability	
Assessor 1, Assessor 2 (Test 1)	1.000 , 1.000	(1.000 to 1.000) , (1.000 to 1.000)
Assessor 1, Assessor 2 (Test 2)	0.565 , 1.000	(0.384 to 0.746) , (1.000 to 1.000)
Assessor 1, Assessor 2 (Test 3)	0.667 , 0.643	(0.461 to 0.873) , (0.425 to 0.861)

14.6.3 Feedback from assessors

Verbal and written feedback was obtained from the assessors. The main themes which were described included: the presence of ambiguous or confusing terminology in the instruction sheets, confusion over whether they were to assess left and/or right feet of participants, and their awareness of recall bias in the repeated trials.

14.7 Discussion

This study was conducted to assess the feasibility of an investigation into the psychometric properties of the three tests within FootFAST, and to inform a power analysis for determining the participant numbers required for a full study.

14.7.1 Main findings

The current (underpowered) study found that the three tests comprising FootFAST demonstrated satisfactory intra-rater reliability and criterion based validity: Kappa scores ranged from 'fair' to 'almost perfect' for both these assessments. In contrast, all three Tests had poor inter-rater reliability across both categorical and continuous results, at least as assessed with the current protocol. Kappa scores ranged from -0.026 to 0.199 (Table 14-3 and Table 14-4). The finding of 'poor' inter-rater reliability affects the feasibility of the main data

collection, as it implies that further investigation is not worthwhile. However, an important issue of note was identified, as discussed below.

14.7.2 Assessors

The assessors chosen for the current study were clinically trained physiotherapists, all of whom were experienced clinicians (a minimum of three years of professional practise); none had used FootFAST previously. This choice of assessor was based on a previous finding that trained clinicians produced more reliable results based on their experience (207). However, the *nature* of the clinical experience was not addressed for the current study, as it was not foreseen as being an influential factor. However, Table 14-7 indicates that the musculoskeletal therapists did demonstrate greater inter-rater reliability. It is possible that this is due to greater experience of these therapists in using musculoskeletal assessments, including exposure to assessments similar to FootFAST, which is based upon visual assessments.

The data for those with musculoskeletal experience indicates more promising inter-rater reliability of all three Tests (ranging from 'moderate' to 'almost perfect'). Musculoskeletal therapists are also anticipated to be the population of choice as regards generalizability of these results (along with sports medics and New Zealand army medics, as discussed below). This is particularly relevant if, as the results indicate, they do not require training to achieve satisfactory results.

14.7.3 Participants

Finally, the group of participants selected for the current feasibility study were a random sample of volunteers from the New Zealand army. The participants for the main study will also be randomly selected from the same group. A recent demographical study identified that New Zealand army soldiers were made up of around 30-40% Maori and Pacific Islanders (80)

(based on verbally provided affiliation). Table 14-1 indicates that the participants in the current feasibility study were made up of 60% Maori and Pacific Islanders (also based on verbally provided affiliation). The anthropometric differences of feet between Maori/Pacific Islanders and New Zealand Europeans (Pakeha) has been documented (80), yet it is unknown whether these differences would have had an effect on the assessors ability to perform the FootFAST protocol. Thus, whether the ethnicity of the participants may have affected the results remains speculative. This will be discussed further in the interpretation of the main study results (Chapter 7).

14.7.4 Feedback from assessors

There were some additional factors raised during the feasibility study which informed the refinement of the methodology design for the main study. In the first instance, feedback from the assessors was helpful in identifying and removing jargon and ambiguities from the printed instructions. For example, based upon feedback of confusion about use of the visual analogue scale, at the top of each new scoring sheet utilised in the main part of data collection, the following statement was added:

"You can place your mark absolutely anywhere along the line".

Another example was the use of the term 'toeing out': this was found to be confusing for some assessors based on their individual interpretation of the phrase. The final information and data sheets were amended, and areas which were changed are highlighted in yellow in the scoring sheets utilised in the main study, shown in Appendix 8.

Further factors which might potentially affect variability of results were also identified. The foot (left or right) which is used during Tests 1-3 needed to be kept consistent. The specific foot which was to be assessed during the current tests was not kept consistent; in some cases the foot assessed during Test1 was not the same as the foot assessed during Test 2 and 3 for a participant. It is potentially possible that the left and right foot of a person are

different both structurally and functionally, which may have a negative impact on the results. This was possibly more an issue of the assessors not reading their instructions for assessing Test 1 carefully enough. To make the instruction more apparent, it has been noted at the top of each scoring chart which foot should be assessed (refer to Appendix 8).

It was noted from feedback from the assessors that due to a small number of participants (N=10) and high number of repeated tests (N=3), they were recalling the previous assessment made for several participants. In order to address this issue, and further increase the quality of the main data collection through the reduction of bias (157), the following change was applied to the methodology: images and videos supplied to assessors will be in a random sequence, which will not be consistent between assessors, or between repeated trials.

Thus, results from the feasibility study highlighted two important areas for refinement in the design of the main study: clarification of the wording of the instruction sheet for assessors, and the inclusion criteria for assessors in terms of clinical experience.

14.8 Further implications for the main study

14.8.1 Assessor experience

If we are to consider the type of assessor who will be delivering a protocol such as FootFAST in the clinical setting, it is logical and ecologically valid to select musculoskeletal physiotherapists as a representative assessor group. The purpose of the current feasibility study and the main study is to assess the reliability and criterion based validity of the three tests, among the practitioner groups who would typically use them. While excluding physiotherapists from the main study who are not musculoskeletal specialists may reduce the generalizability of the results to a wider professional group, on the other hand, their exclusion would increase the internal validity of the main study (Chapter 7). At this stage, it was deemed more appropriate to explore the inter and intra-rater reliability, and criterion based

validity of FootFAST within the groups of clinicians who are most likely to encounter such a test. If results support the use of the tool, then future studies could assess the performance of the tool when used by other groups of clinicians.

As results of the main study will be generalised to other groups of clinicians likely to use FootFAST, it was considered important, in hindsight of the completion of the feasibility study, to include other representative groups as assessors for the main study, including sports medic practitioners and NZ army medical staff (medic). Sports medics are anecdotally the most common choice of sports teams and coaches for performing preseason injury screening for financial, practical, and feasibility reasons (218). Similarly, NZ army medical staff (termed 'medics') completes screening protocols for injury risk for every new solider entering basic training. In keeping with the rationale outlined above for the inclusion of musculoskeletal physiotherapists as assessors, it was considered that the research would be somewhat incomplete if it did not include those practitioners who would typically be using such a protocol. If such groups were not included, it is unknown the extent to which the results of the study would be generalizable to them.

Furthermore, the screening protocol FootFAST was developed with groups such as the NZ army in mind: this includes considerations of time for completion, level of training required, equipment requirements, and FootFAST's potential sensitivity to identify risks of specific injuries of the lower limb. Thus, the proposed future application of the screening protocol is the justification of including NZ army medics in the assessor group.

14.8.2 Sample size estimation

One of the aims of the current feasibility study was to calculate the sample sizes required in order to attain a minimum statistical power level of 80% in order to avoid type II error (β). It is important to ensure that a study has adequate power, because with the results of any scientific investigation, it is possible that an association seen may have occurred due to

chance. It is also possible that an association may be missed due to chance. Study numbers should be big enough to minimise the occurrence of either. When making the assessment of how large a study should be in order to ensure it can detect an association if one exists, it is referred to as powering a study. The power of a study is measured from zero to one, or from 0 to 100%, and the score represents the probability that the study could detect an association if one exists in the general population. Eighty per cent power is typically regarded as the acceptable level, which means: if an association truly exists in a population, there is an 80% chance that the study will show this. (157)

The number of participants and assessors required for the main study was calculated using the results presented from the musculoskeletal physiotherapists, given the superior results from this group, and as this is the only sub group from the feasibility study who will be included in the main study (with the addition of sports medics and NZ army medics). To ensure a statistical power of 80% (219), and therefore a Kappa 95% confidence interval of +/- 0.20 (or less), a minimum of eighteen assessors and eighteen participants would be required, presuming the scores allocated will be normally distributed ((219); p.444 Figure 2).

14.9 Limitations

The most obvious limitation of this feasibility study is that it was not possible to ensure that the assessor group included the different types of clinicians of interest. This may limit the accuracy of the sample size requirements, depending on how the results from sports medicine practitioners and NZ army medics compare to those from physiotherapists. As no previous studies have included such a diverse range of assessors, it is impossible to determine definitely how the sample size requirements would be affected by this.

The feasibility study is also limited in that there will be no opportunity to check any of the amendments made prior to the main part of data collection, i.e. performing a second feasibility study. Although all the amendments made have been justified, it is not clear that

all of them will be optimally effective. However, this will always be an issue in the use of feasibility studies (203).

14.10 Conclusions

The current feasibility study had the aim of answering the research questions:

1. Is an intra-rater reliability assessment of the three FootFAST tests feasible?
2. Is an inter-rater reliability assessment of the three FootFAST tests feasible?
3. Is a criterion based validity assessment of the three FootFAST tests feasible?
4. What is the size of study required in order to produce confidence intervals of K<0.2 with 80% power?

Based on the results of the feasibility study, and the discussion, an investigation into the inter and intra-rater reliability, and criterion based validity of the three FootFAST tests is feasible. Further to this, the size of study needed to produce confidence intervals of Kappa<0.2 was calculated to require a minimum of eighteen participants and eighteen assessors.

The highly variable results obtained from both the intra-rater reliability and criterion based validity assessments are likely due to the low numbers in both the assessor and participant groups. The increased numbers of both assessors and participants in the main study was expected to address this limitation. The somewhat poor results obtained from inter-rater reliability assessments could be due to the choice of assessor, an issue which has been addressed for the main data collection study, or because the study was underpowered, or may be due to other, yet unknown reasons. The results from the main study will provide further clarification on this.

In the instances where acceptable levels of agreement were attained (Kappa has been defined as being clinically acceptable with values from Kappa>0.4 (217) to Kappa >0.7 (220)) it was hypothesised to be due to clear instructions, simple measurement scales, and

relevance to the assessor (as noted in previous studies) (67, 153). This knowledge was used when re-wording the assessor instruction sheets for the main study (Appendix 8).

The issues identified in the methodology during the feasibility trial were addressed, as discussed above. All changes made to the information sheets provided to assessors are highlighted in Appendix 8.

14.11 In the next Chapter

The current Chapter presented the feasibility study which was required to inform the main stage of data collection. The next Chapter presents the main data collection of the thesis, titled: *"What are the psychometric properties of the injury risk assessment tool: FootFAST?"* It follows the same outline as the current Chapter.

15 Chapter 7. What are the psychometric properties of the injury risk assessment tool: FootFAST? A repeated measures cohort study in a military population

15.1 Overview of Chapter

Following on from the feasibility study presented in the previous Chapter, this Chapter presents the main research study which aimed to answer the final three research questions of the current thesis (refer to Chapter 1, p.14)

Further to these, based on the findings from the feasibility study, the current study aimed to address an additional question, which was:

Can New Zealand army medics, sports medics and musculoskeletal physiotherapists use FootFAST to a comparable level of reliability?

This Chapter is the last stage in answering the research questions above, and is a focal part of the current thesis. Essentially, this is an investigation of whether the three tests within FootFAST are justified for continued use in a clinical setting. The importance of this knowledge is therefore twofold; firstly to ensure that clinical practise is justified, in terms of the continued use of the FootFAST tool in its current form, and secondly to ensure that the continued development of the FootFAST test as an injury prevention protocol is justified.

The study followed a repeated measures cross-sectional design (157): eighteen assessors evaluated eighteen participants on three separate occasions (numbers included were based on a power calculation using the results from the feasibility study (Chapter 6)). The eighteen assessors comprised three groups: sports medic practitioners (N=6), musculoskeletal physiotherapists (N=6), and New Zealand army medics (N=6). The reasoning behind using such a mixed assessor group was twofold: to investigate whether the reliability of the test was dependant on the group delivering the assessment, and also to ensure that the protocol could

be applied to an appropriate standard by medics within the NZ army. This has been discussed previously (refer to Chapters 3, p.63-68 and Chapter 6, p.130-134).

The methods, which were essentially the same as those utilised in the feasibility study (subject to the amendments made during the feasibility study, as presented in the discussion section of Chapter 6), are presented again in brief followed by results and discussion. Any deviations from the methods used in the feasibility study (Chapter 6) are highlighted in the Appendices and presented in detail.

15.2 Methods

This is a repeated measures cross-sectional study (157) with a participant sample of convenience recruited from Burnham Military Camp, Christchurch, New Zealand between the 29th September 2012 and 2nd October 2012. Ethical approval for the study was granted by the University of Otago Human Ethics Committee, and Maori consultation was undertaken (refer to Appendix 9). Lt Col Andrew Dunn from the New Zealand Medical Corps provided consent for access to the soldiers included as participants.

15.2.1 Participants

Participants consisted of eighteen male New Zealand Army soldiers, aged between nineteen and twenty five years. All participants met the inclusion criteria (as detailed below). All participants were volunteers and had signed the consent form (refer to Appendix 6).

15.2.2 Inclusion criteria

Participants were selected from a group of New Zealand army soldiers from 2nd/1st Battalion Infantry Regiment who were present for first parade between 29th September 2012 and 2nd October 2012, at Burnham Camp, Christchurch, New Zealand. The first eighteen soldiers to volunteer who met the inclusion criteria were selected.

15.2.3 Exclusion criteria

Participants were excluded if they had a current injury, or one in the previous three months that affected their ability to walk.

15.2.4 Assessors

There were eighteen assessors in total, recruited from three groups of practitioners. The first six assessors consisted of clinically trained (Bachelor of Physiotherapy), clinically experienced (minimum of three years), musculoskeletal specialised physiotherapists from the School of Physiotherapy, University of Otago. The second six assessors were sports medic practitioners who were registered with Sports Medicine New Zealand (SMNZ), and thus met the SMNZ criteria, for a minimum preceding period of three years (such information from SMNZ was provided by the respondents). The final six assessors were New Zealand army medics who had graduated from the Defence Force Health School (DFHS) and had at least three years of experience within the Defence Force Medical Centres (DFMC). The medics also had to have been posted in Burnham Military Camp, Christchurch, New Zealand during October 2012 in order to receive the recruitment email.

A recruiting email (refer to Appendix 5) was distributed by the researcher among all the members of the groups detailed, and the first six voluntary respondents who met the inclusion criteria were chosen for the study. For the musculoskeletal physiotherapists, the email was distributed among the staff email register. The same recruiting email was sent to all sports medics who were currently practising in Otago (information obtained from the Sports Medicine New Zealand website (http://www.sportsmedicine.co.nz/smnz_register)). Finally, for the NZ army staff, the email was distributed to all medics who were currently posted to Burnham Military Camp (information obtained from the Burnham Medical Treatment Centre staff). All eighteen assessors had signed the consent form (refer to Appendix 7).

15.2.5 Inclusion criteria

To be included as an assessor in this study, the assessor had to meet one of the following criteria:

1. Working for the University of Otago School of Physiotherapy Clinics during October 2012. The assessor had to have had a minimum of three years of clinical experience as a musculoskeletal physiotherapist.
2. Registered with SMNZ currently in October 2012, and for the preceding three years. The assessor had to have had a minimum of three years of clinical experience as a sports medic practitioner.
3. Working as a medic for the NZ army at Burnham Military Camp, Christchurch, New Zealand during October 2012. The assessor had to have graduated from DFHS, and had a minimum of three years of clinical experience working for DFMC as a medic in the New Zealand army.

15.2.6 Exclusion criteria

Assessors were excluded if they had previous experience using FootFAST.

15.2.7 Equipment

Two high speed digital video cameras (Sony Alpha A37 DSLR / SLT; Sony Limited, New Zealand) were used to capture still images of the hindfoot, and videos of both single leg stance and heel raise performances. Such images and videos were used by assessors for later analysis. The RS Scan™ mat (RS Scan International, Belgium) and accompanying Gait™ software were used to capture the centre of pressure data and underfoot pressure characteristics of the foot during single limb stance and heel raise performances. The Silicon Coach™ (SiliconCOACH, New Zealand) software programme was used to calculate hindfoot angle. A black permanent marker was used to outline anatomical reference points described

below. One 1.5 metre measuring tape was used to calibrate the camera and measure the distance between the ischial tuberosities. Finally, nine reflective markers, 10mm in diameter, were used to mark reference points for anatomical angular measurement. Specifically, these were placed bilaterally on each malleolus, the medial condyle of the tibia, the lateral aspect of the head of the fibula, also on the palpable tip of the calcaneus, and a marker was placed on the Achilles tendon at the fusion of the gastrocnemius muscles with the triceps surae.

15.2.8 Procedure

There were three parts to the procedure, corresponding with the three tests within FootFAST. For the first Test, the environment was set up with a camera placed as far from the viewing field as possible to minimise for parallax error, and zoomed in so that minimal excess picture was included (208). For the second and third Tests, the environment was set up as follows: the RS Scan™ was placed in the centre of the room so that the participant could not fall onto any other piece of equipment, or use a wall for support. Two cameras were set up to record images either from the medial side and from the front (Test 2) or from the lateral side and from the rear (Test 3) of the planted foot. These cameras were situated as far from the mat as possible (12.5 meters away), and perpendicular to the mat to avoid parallax error. A right angled triangle was created in order to ensure that the camera was perpendicular to the mat, with sides measuring 3m / 4m / 5m (according to Pythagoras' Theorem). The foot and the lower half of the shank were included in the field of view.

15.2.9 Test 1 (Hindfoot angle)

The participant was in bare feet and wore clothing above knee level for this trial. Two separate digital images were taken, one with joint markers attached- for reference points when measuring the angles within Silicon Coach™ (SiliconCOACH, New Zealand), and one without for visual assessments. Silicon Coach™ (SiliconCOACH, New Zealand) has beenA

distance reference line 10cm in length was drawn on the wall, within the field of camera view and perpendicular to the floor. This distance was used for scalar calibration of these images using CAD software program Silicon Coach^{TM} (this is in accordance with the calibration process outlined in the software's reference manual - SiliconCOACH, New Zealand).

The distance between the ischial tuberosities (identified through palpation) was measured using the 1.5 metre measuring tape; this was then apportioned as the distance that the feet were placed apart. Lines were drawn on the ground, perpendicular to the wall, to indicate the distance apart as measured between the ischial tuberosities. The participant placed both feet on these lines, with the line running through the middle of their heel and under the second toe. Although this may not be indicative of typical stance, it minimises parallax error as the heel is optimally aligned with the optical field of view. (209, 210)

The lower limb from above the knee joint distally was included in the field of view and a digital image was captured (as shown in Figure 15). Once an initial image was captured, joint markers were placed on the participant. These markers were placed in similar locations as used by Davison et al (35), as outlined above. Marks which were 2cm in length were drawn with a permanent marker on both lateral and medial borders of the calcaneum where it made contact with the ground. The participant was returned to the same position as for the first image, and a second image was captured. (Hindfoot angle).

15.2.10 Test 2 (Single leg stance)

All marks and markers were removed from the participant. The participant then stood on the RS Scan^{TM} mat with the axis of the foot parallel with the longest side of the mat, and was permitted to position their upper body and arms in their preferred position. When instructed to, the participant lifted one foot off the ground, closed their eyes and attempted to remain balanced for as long as possible. Once the participant could no longer maintain balance the trial was accepted as being complete. There were no practice attempts permitted and no

repeated trials, as this is the procedure used within FootFAST (1). The ankle joint was the focal point of the footage obtained in this performance, as it has been shown to be the joint of primary importance during single leg stance (211).

15.2.11 Test 3 (Heel raise)

Participants again stood on the RS Scan™ mat and were permitted to hold their upper body and arms in their preferred position. When instructed to, participants lifted one foot off the ground and attempted to rise unto the toes of the planted foot. The participant was asked to maintain this posture for about three seconds. There was only one attempt allowed for this test and no repeated trials, as this is the procedure used within FootFAST.

15.3 Data collection and analysis

The images captured for Test 1 without bodily markers, along with the video footage captured for Tests 2 and 3, were edited to ensure minimal unnecessary footage or background images were present. These were then included in a DVD, and sent to the eighteen assessors. The order of videos was randomised for each assessor, and a different order sequence was used for each repeated test; randomisation was achieved using a random sequence generator within Microsoft Excel™. Assessors used this visual information to grade the participants using the FootFAST procedure. All written information provided to the assessors is included in Appendix 8, and changes made from the feasibility study are highlighted.

For each of the data items below, the calculations performed were identical to those described during the feasibility study, with the exception of the participant (N=18) and assessor (N=18) numbers. To review these calculations, refer to the sections of the thesis as indicated.

Based on the discussion presented in Chapter 6; Data collection and analysis; Categorical and continuous scales, the results were again analysed as both categorical and continuous data where appropriate.

15.3.1 Inter-rater reliability

Refer to Chapter 6; Methods; Data collection and analysis; Inter-rater reliability.

15.3.2 Intra-rater reliability

Refer to Chapter 6; Methods; Data collection and analysis; Intra-rater reliability.

15.3.3 Criterion based validity

Refer to Chapter 6; Methods; Data collection and analysis; Criterion based validity.

15.4 Statistical calculations

All statistical analyses were conducted using STATA 7.0 software (215). Dr Andrew Gray (Biostatistician, School of Preventive and Social Medicine, Dunedin School of Medicine) was consulted with regards to choice of appropriate statistical test, and confirmation of results, with the exception of the calculation of assessor-group difference between means, as outlined below.

15.4.1 Inter-rater reliability

Refer to Chapter 6; Methods; Statistical calculations; Inter-rater reliability.

15.4.2 Intra-rater reliability

To calculate intra-rater reliability, the results given on the three separate occasions of testing were compared using Cohen's weighted Kappa. Specifically, the agreement level for all three assessments (Tests 1, 2, and 3) of each patient individually was calculated, then averaged across all participants, for all assessors.

It was first determined if the assessors in the three practitioner groups (sports medics, physiotherapists, and NZ army medics) produced similar scores of intra-rater reliability. The aim of this process was to investigate whether assessments and corresponding variances were consistent and if not, where the variance lay. If the assessments were consistent then the results would be pooled.

To do this, the following methodology was used (advised by Dr Josie Athens (Biostatistician, School of Preventive and Social Medicine, Dunedin School of Medicine)):

1. An intra-rater reliability score from each assessor individually was calculated in accordance with the example provided above (refer to p.119).
2. This provided the average intra-rater reliability score for each of the eighteen assessors, within each of the three tests.
3. The reliability scores from the six assessors within each of the three groups of practitioners (musculoskeletal physiotherapists, sports medics and NZ army medics) were averaged. This calculation was performed for each test separately.
4. The averaged reliability scores gained for each of the three assessor groups were assessed for differences using a P-value obtained from a T-test for differences between means. If there were significant differences between the mean scores of assessors from each group, their scores would not have been considered collectively. This was performed for each test separately: for example, average intra-rater reliability of musculoskeletal physiotherapists *versus* average intra-rater reliability of sports medics, for Test1 (Hindfoot angle). P-values from these comparisons were obtained from a test of difference between means.
5. The P-values were averaged across all three tests, for each pair of group comparisons.
6. If the results were not significantly different, they were combined. If not, all three assessor groups would be assessed separately, as individual groups with unequal variances.

7. The intra-rater reliability scores from the assessors in each group, who each performed the assessments three times, were averaged for each test individually. (Table 15-4)

15.4.3 Criterion based validity

Refer to Chapter 6; Methods; Data collection and analysis; Criterion based validity.

15.5 Results

All eighteen assessors completed their requirements for the study. Results were received from assessors between 3rd December 2012 and 16th March 2013. Four assessors returned incomplete results: three did not perform the repeated tests, and one did not perform the required number of repeats (only one, when two were required). These assessors were contacted and requested to complete their result sets. All result sets were consequentially received in their completed form.

15.5.1 Inter-rater reliability

Table 15-1 15-1 shows that across all three FootFAST tests, the inter-rater reliability was 'moderate', and reached a clinically acceptable level of agreement (K>0.4). The confidence intervals for agreement in all three tests had a width of approximately 0.100.

Table 15-1: Fliess' weighted Kappa for inter-rater reliability of results assigned categorically

Test	1 (Hindfoot angle)	2 (Single Leg Stance)	3 (Heel Raise)
Fleiss's weighted kappa (width of 95% confidence interval)	0.450 (0.399 to 0.502)	0.450 (0.418 to 0.481)	0.450 (0.400 to 0.499)

The Fliess' weighted Kappa (FWK) calculated for continuous results in each of the three tests within FootFAST was 'fair' (

Table 15-2 15-2), and below the clinically acceptable level of agreement. Again, confidence intervals for agreement in all three tests were approximately 0.100.

Table 15-2: Fliess' weighted Kappa for inter-rater reliability of continuous results

Test	1 (Hindfoot angle)	2 (Single Leg Stance)	3 (Heel Raise)
Fleiss's weighted kappa (width of 95% confidence interval)	0.341 (0.298 to 0.384)	0.353 (0.313 to 0.392)	0.356 (0.290 to 0.422)

Results presented in Table 15-3 show that for all three tests, and across all eighteen participants, the agreement levels were varied- all three tests had agreement levels from ‘fair’ to ‘almost perfect’.

Table 15-3: Flies weighted Kappa scores for inter-rater reliability, by participant, for each test

Participant	1 (Hindfoot angle)	2 (Single Leg Stance)	3 (Heel Raise)
1	0.41	0.46	0.46
2	0.43	0.68	0.48
3	0.36	0.28	0.27
4	0.23	0.41	0.31
5	0.31	0.68	0.38
6	0.46	0.50	0.58
7	0.43	0.56	0.58
8	0.47	0.49	0.77
9	0.30	0.36	0.33
10	0.39	0.35	0.28
11	0.49	0.37	0.39
12	0.41	0.36	0.41
13	0.31	0.48	1.00
14	0.60	0.43	0.23
15	0.34	0.41	0.39
16	0.76	0.35	0.28
17	0.77	0.68	0.51
18	0.60	0.28	0.66

15.5.2 Intra-rater reliability

Table 15-4 shows the P-values obtained from a T-test of differences between means (i.e. the average grade assigned by the assessor groups). This was a comparison of the reliability scores between the three groups. The P-values indicate that there was no significant difference between means of the three groups of assessors with respect to their intra-rater reliability scores.

Table 15-4: Comparison of intra-rater reliability results attained between assessor groups.

Group comparison	P-Value (95% confidence interval)
Musculoskeletal physiotherapists vs. NZ army medics	0.667 (0.505 to 0.723)
Musculoskeletal physiotherapists vs. Sports medic practitioners	0.523 (0.425 to 0.754)
Sports medic practitioners vs. NZ army medics	0.898 (0.656 to 0.934)

For all three tests within FootFAST, there was an 'almost perfect' level of intra-rater agreement (Table 15-5). An intra-rater reliability score was calculated for each assessor individually, then averaged for the assessor group (N=18). The level of agreement was consistent across all three groups of raters with no significant differences between means of the groups (Table 15-4). In general, Test 3 produced the most variable results based upon the confidence interval widths. However, it also produced the highest overall level of agreement. Test 2 produced the most consistent scores, as it shows the smallest confidence interval width.

Table 15-5: Cohen's weighted Kappa calculations for intra-rater reliability: averaged intra-rater results for all assessors (N=18), for each test.

FootFAST test number	CWK level of agreement (95% confidence interval width)
1: Hindfoot angle	0.920 (0.881 to 0.958)
2: Single legged stance	0.890 (0.839 to 0.941)
3: Heel raise	0.940 (0.860 to 1.00)

15.5.3 Criterion based validity

The results in Table 15-6 show that the criterion based validity assessment for Test 1 has a 'moderate' agreement (based on Cohen's weighted Kappa) when compared to Silicon Coach™ angular measurements, and thus met the minimum standard for a clinically acceptable level. For Test 2, there was an 'almost perfect' agreement when compared to COP tracking. Finally, for Test 3, the FootFAST test had 'fair' agreement when compared to COP tracking, but did not achieve a clinically acceptable level.

Table 15-6: Cohen's weighted kappa (CWK) calculations for the three tests within FootFAST, calculating criterion based validity (comparison of assessor results with Silicon Coach™ (Test 1) and Gait™ (Tests 2 and 3) software results)

Test	1 (Hindfoot angle)	2 (Single Leg Stance)	3 (Heel Raise)
CWK (width of 95% confidence interval)	0.448 (0.375 to 0.521)	0.830 (0.750 to 0.911)	0.390 (0.305 to 0.445)

15.6 Discussion

This study aimed to answer the final three research questions of the current thesis which were:

3. What is the inter-rater reliability of the three tests within the FootFAST screening protocol?

4. What is the intra-rater reliability of the three tests within the FootFAST screening protocol?

5. What is the criterion based validity of the three tests within the FootFAST screening protocol?

During the feasibility study it was identified that eighteen assessors and eighteen participants were required in order to attain a minimum statistical power level of 80% and thus avoid type II error (β). For the main study, a full data set was obtained from each of the eighteen assessors and therefore bias is minimised in this respect. Results indicated that all three tests

had acceptable levels of both inter and intra-rater reliability, and two (Tests 1 and 2) out of three tests had acceptable criterion based validity.

15.6.1 Summary of key findings

The assessment of inter-rater reliability found that all three tests within FootFAST were at a clinically acceptable level, with agreement of Kappa>0.4 (216). For each of the three tests, Test 1 (hindfoot angle), Test 2 (single legged stance), and Test 3 (heel raise) the agreement was consistently K = 0.450 (

Table 15-1 15-1). The confidence intervals measured were all narrow, less than 0.20 in width, and for two tests the lower limits were also above the clinically acceptable level of agreement, again increasing the confidence of the results; the single exception was the Test 1 lower limit, which presented a 95% CI of 0.399-0.501.

The investigation of inter-rater reliability for continuously assigned outcomes produced Kappa scores that were below the clinically acceptable level of agreement across all three tests (for Tests 1, 2 and 3 these were K=0.341, 0.353 and 0.356 respectively (

Table 15-2 15-2)). In addition, the application of statistical methods to both the categorical and continuous data sets were equally uncomplicated. Thus, given the knowledge that the categorical measurement scale represents a classification system more conducive to better reliability outcomes, it would seem that the existing categorical method of assigning results is optimal for the FootFAST protocol. This is, of course, a positive finding as it is currently the classification scale used within FootFAST and most other visual assessments of function (14).

The main reason for the differences between categorical outcome agreement and continuous outcome agreement is likely to be due to choice. With increasing options, there will be increasing variability of response (157). Further to this, as most functional assessments used by clinicians require a response on a categorical scale (14), the reduced

familiarity of using a continuous scale may have contributed to the observed variability. As FootFAST is designed to produce results on a categorical scale, the discussion will primarily focus on the relevant categorical outcomes.

For the assessment of intra-rater reliability, it was found that all three tests within FootFAST had almost perfect levels of Kappa agreement (K=0.920, 0.890, and 0.940 respectively (

Table 15-5 15-5)). In addition, the confidence intervals were small- less than 0.20 in width. There were no significant differences in the reliability scores between the groups of assessors (p-values greater than 0.523 for each group comparison (

Table 15-4 15-4)). The significance of this finding is discussed further below. Criterion based validity assessments of the tests within FootFAST produced clinically acceptable levels of agreement for two out of three tests (
Table 15-3 15-6). The exception, Test 3, had a marginal Kappa score (Kappa = 0.390) that approached clinical significance (Kappa >0.4). Test 2 had the highest level of agreement for criterion based validity, which was found to be 'almost perfect' (Kappa=0.92).

15.6.2 Inter-rater reliability; effect of participants

Table 15-3 shows the results of an investigation into the level of inter-rater agreements for each participant was conducted. This was an attempt to identify whether the individual participant (or in ecological settings, the patient), had a confounding effect on the ability of assessors to produce consistent results. Taking the definition of a confounding variable as being that which interferes with the observation we are interested in (157), in the current example, potential confounding variables include the anatomical structure of the participants' foot and ankle, and subcutaneous fat in the same areas. Both of these could potentially inhibit the ability of the assessor when performing the FootFAST protocol. For example, if there

were a large amount of subcutaneous fat around the ankle joint and achillies tendon, it may be difficult for the assessor to identify the alignment of the hindfoot as bony landmarks may be obscured. This suggestion was speculation, as there were no previous research studies found that included such an investigation in their data analysis; however, it seemed important to include the potential effect of participants in the current study in order to account for all potential sources of confounding.

For all three tests within FootFAST, there was a highly varied level of agreement between assessors for each of the participants, and no apparent trend based on the participants as individuals (

Table 15-6 15-6). Thus, it is unlikely that the participants produced a confounding effect on the results. This finding is important as it indicates that the ability of the practitioner to assess a patient is not dependant on the individual. It further suggests that the variation seen between the assessors is related to their own interpretations rather than to variability in the patients.

15.6.3 Inter-rater reliability

Inter-rater reliability for all three tests was assessed from both categorical and continuously assigned outcomes. The categorical inter-rater reliability outcome is more relevant as the tests are designed to be assessed on a categorical scale identical to the ones utilised in the current research study. The purpose of conducting the continuously assigned outcome analysis was to investigate whether it would produce better reliability scores, or was more appropriate to apply statistical methods to, as discussed in Chapter 6. If this was found, then the possibility of changing the required outcome from the tests would be explored.

All three of the tests within FootFAST reached a clinically acceptable level of inter-rater reliability. Reasons why these tests reached the clinically acceptable level of agreement (overall) are thought to be:

1. The comparative simplicity of the response required from the assessors (67, 138, 146, 149); in the current example, assessors were required to assign a participant into one of only five groups.
2. The quality of the images presented (67, 149). Care was taken to produce high quality images of the three tests, and all footage from all participants was exactly the same; including field of view, lighting, and proximity to the camera.
3. The use of a feasibility trial, which helped refine the protocol including instructions or wording which was potentially confusing and/or ambiguous to the assessors (137).
4. The inclusion of practitioners who had a minimum of three years of clinical experience in a relevant specialised area (68, 138).

15.6.3.1.1 Test 1 (hindfoot angle)

The inter-rater reliability for the categorical outcomes was 'acceptable' (Kappa=0.450 (95% CI 0.399 to 0.502)) for Test 1. Taking the clinically acceptable minimum value for agreement as being Kappa greater than 0.4 (216), the overall level of agreement for this test meets the clinically acceptable level, although the lower confidence interval limit falls just outside the cut-off value of 0.4. The implication of this finding is that practitioners can continue to use this test within the clinical setting, knowing that it has an acceptable level of inter-rater reliability and thus is justified.

It was found that several previous studies had investigated the inter-rater reliability of a visual assessment of hindfoot angle (68, 221-223). Findings from all previous studies were consistent with the findings here. Three previous studies were less comparable to the current study, and are summarised in brief (as follows), whereas the Haight (68) study was highly comparable and is discussed in detail.

Watson and MacDonncha (223) attempted to define the reliability for repeated assessments of the hindfoot, when a visual estimate of the exact angle was made. They found

that the 95% confidence interval for repeated measurements were 0.98^0 to 2.5^0, and thus the visual estimates were reported as having acceptable inter-rater reliability. Similarly, Thomson (221) attempted to define the reliability of the visual estimate of the hindfoot angle. They identified a confidence interval of $_{+}/^{-}2^0$, and concluded that the visual estimate had an acceptable level of inter-rater reliability. Pownall et al (222) investigated the inter-rater reliability of angular estimates of patients Achilles/calcaneus and ankle/fibula alignments. They reported ICC>0.92, and therefore concluded that the visual estimates had excellent reliability.

Haight (68) also investigated the inter-rater reliability of visually based hindfoot angle assessment. This study was highly comparable to the methodology of the current study. Two blinded assessors completed a right and left foot assessment on eighteen participants (University students, described as 19-24 year old 'active' individuals, thus are comparable to the participants in the current study). Inter-rater reliability of their measurements was reported as 'satisfactory', with intra-class correlation coefficients (ICC) of 0.56 to 0.65. Their outcome is similar to the agreement reported in the current study: a Kappa value that was 'acceptable'. However, there were a number of differences between the two studies which may have an effect on the comparability of the results.

Firstly, the visual assessment in the Haight (68) study required assessors to estimate an exact angle for the hindfoot (thus, only a continuous scale was utilised). In the current study, the continuously assigned results produced a Kappa=0.341 for the assessment of Test 1. Although FootFAST does not require a response on a continuous scale, it is useful to compare the results gained here to that of the previous study. In this respect, Test 1 produced a lower inter-rater reliability than the study performed by Haight (68). A likely reason for this is that the current study had sufficient numbers of assessors to produce power of 80%. Thus, it is possible that the Haight (68) study produced inflated results due to the higher variability

present in underpowered studies, and the decreased likelihood of an unpowered study being able to detect a real association (157).

Haight's (68) use of ICC as a comparison method is problematic. Performing an intra-class correlation calculation for the purposes of assessing inter-rater reliability is not ideal (151). The reason for this is a correlation that is measured to be perfect (ICC of 1.0) may still not include a single pair of results which are the same. The Kappa statistic in its various forms is much more appropriate in this instance (151, 216), as the primary outcome of interest are results which are in *agreement*, and not which *correlate*. Several studies referred to in the following discussion also utilised ICC values. The authors of studies mentioned which utilised ICC were contacted to provide the original data, so that Kappa statistics could be calculated, but no response was received.

Additional differences with Haight (68) include how participants were instructed to stand before performing the test. The stance position of participants in the current study was highly standardised. Participants were required to place their foot along marked lines, and the distance apart was consistently the same- the measured distance between their ischial tuberosities. The previous study asked participants to walk ten paces then stop, in order to produce a 'natural' stance (68). It is unknown what effect these differences may have had on the observed reliability results. The problem of trying to determine which stance position is most appropriate has been discussed by previous researchers (159). However, there is still no consensus on this issue, which needs to be addressed in future studies.

As the effect potential differences between stance positions has on the comparability of the studies is unknown, it was assumed that the results were comparable. Thus, to summarise, two studies to date have investigated the inter-rater reliability of hindfoot angle measured visually; both of which were performed to an acceptable level of methodological rigor. Both the current and previous study, despite their differences, found that inter-rater reliability

reached a clinically acceptable level. Therefore, the present results suggest that this test is suitable for continued use within the clinical setting.

15.6.3.1.2 Test 2 (Single leg stance)

Inter-rater reliability for the categorical results was 'acceptable' Kappa=0.450 (95% CI 0.418 to 0.481) for Test 2. Taking the clinically acceptable minimum value for agreement as being Kappa greater than 0.4 (216), this test met the standard and a narrow 95% confidence interval gives a high level of confidence in the results.

Previous studies by Trojian (153), Haupstein and Goldie (58), and Finoff et al (145), also investigated the inter-rater reliability of the single leg balance test using video footage of participants' performance. Two of these studies were identified as being of high quality in the literature review, and are presented in detail in Chapter 3. The study by Haupstein and Goldie (58) (as discussed previously), found the inter-rater reliability of their balance test to be high: ICC=0.81 (14 assessors, 20 participants). Trojian (146) included only two assessors: one physiotherapist and one sports medic, and forty participants. The inter-rater reliability score reported was almost perfect- Kappa=0.898. Although the current study and the two previous studies identified so far reported agreement levels above the minimal acceptable level for clinical use (216), it is noted that the previous experiments reported agreement levels substantially higher than in the current study.

Several differences between the studies may have contributed to the differing results. Firstly, the Trojian (153) study utilised forty participants, but only two assessors. Including only two assessors would have potentially improved their reliability scores when compared to the current study, as reduced assessor numbers implies an increase in reliability through a reduction in response variability (151). Both assessors did however have more than three years of experience, which was the same as the current study. Despite this, it is unlikely that a study based upon only two assessors and forty participants has sufficient statistical power

(given the information presented in Donner (219)); the difficulty in comparing current results with a study which was not adequately powered has been discussed. Secondly, the previous study only required a dichotomous response; i.e. in two categories: 'fail' or 'pass' (whether balance was maintained for ten seconds or not). Thus, the outcome measure used in the previous study was much more simplistic than the current study, which required classification of participants into one of five categories.

The justification for using only two categories in the Trojian (2006) study was not discussed by the authors. Based on the probability theory, with increasing numbers of categories to choose from, there will be increasing variability of results (157). Thus, when the number of response options increase, there is likely to be reduced agreement. This could potentially have contributed to the reduced level of agreement seen in the current study compared to the previous.

Reducing the number of assessment categories has not been considered as an option for the development of FootFAST. This is for two main reasons: firstly, it is of concern that doing so may decrease its effectiveness as an injury risk screening protocol. The numbers of degrees of freedom within the human bipedal system are countless: the likelihood of being able to adequately discriminate patients into only two categories, and still preserve the presence of different constructs is slim (which is part of discriminative validity), therefore it seems more appropriate to keep the scale broad. Secondly, the use of only two outcome groups has not been justified or discussed in previous studies; it may have been that utilising only two categories was chosen for ease of the research. In summary, it is unknown whether a dichotomous assessment would be sufficient within a clinical setting.

Finnoff et al. (145) investigated the inter-rater reliability of single leg stance, based on a visual measurement scale similar to the protocol utilised in the current study. Their study included thirty participants and three assessors, and was powered to 80% based on the results

of their previous investigations (145). Finnoff et al. (145) identified single leg stance as having an inter-rater reliability ICC of 0.83 (95% CI 0.71 to 0.91), and thus above the minimum value for clinically acceptable reliability. Again, the study of comparison found a higher inter-rater reliability than that seen in the current study.

There are a number of possible reasons for this. Firstly, despite both studies being powered to 80%, the previous study utilised thirty participants, but only three assessors, whereas the current study included eighteen assessors. Secondly, the previous study scored movement aberrations based on the occurrence of 'faults' (for example, the participant moving their hands off their hips), whereas the current study required putting participants into one of five categories based on the amount of observed unsteadiness. It is known that studies using ordinal categorical scales require larger participant samples than those which utilise continuous scales (157). However, it is unclear what the effect of the two different scoring systems may have had on the comparability of the inter-rater reliability of the tests. Finally, as has been discussed, the ICC which was used to calculate agreement is not an optimal statistical method to use for interpretation of reliability (216).

In summary, three studies to date have assessed the inter-rater reliability of the single leg stance test as a visual assessment of balance performance (two of which were powered to 80%). Despite some differences between the current and previous studies, mainly with regard to the reported outcomes, all studies reported the inter-rater reliability for the single leg stance test to be above the clinically acceptable level.

15.6.3.1.3 Test 3 (Heel raise)

The inter-rater reliability for Test 3 was Kappa=0.450 (95% CI 0.400 to 0.499). Taking the clinically acceptable minimum value for agreement as being Kappa greater than 0.4 (216), the overall level of agreement for this test was 'acceptable'. There is confidence in these results given the narrow confidence interval.

A previous study by Dennis et al (60) investigated the inter-reliability of the calf raise test, which is very similar to Test 3. They found that the inter-rater reliability was ICC>0.80 (ten participants, two assessors), which was above their acceptable minimum level.

Beyond this, a previous study by Eechaute (149) investigated the inter-rater reliability of a hop test, and later added to this by investigating the tests evaluative and discriminative ability (144). Although the two movements are different, the functional process of a heel raise is essentially the same movement as is performed in the preparation phases (1-3) of a single leg hop:

1. The participant stands with hands on hips, facing forwards.
2. The participant lifts one leg off the ground and briefly balances on the planted foot.
3. The participant makes a fast and forceful contraction of the gastrocnemius muscle, raising the heel of the planted foot.
4. The participant jumps forwards, covering as much distance as possible, and lands on the same leg as was planted. (149)

In addition, both performances have a similar grading system based on observable 'faults'. Eechaute (149) reported ICC values of over 0.9 for two observers and fifty eight participants. In their further research, the test was found to be discriminative: it could distinguish between ankles with and without a chronic instability, but the evaluative ability of the test was lacking (144). Although all three studies discussed support the inter-rater reliability of these similar tests, it is of note that the previous studies produced higher levels of reliability.

Explanations as to why the previous studies produced a higher level of reliability include that, as indicated above, use of Intra-class Correlation Coefficient (ICC) to assess reliability is not appropriate (151, 216). If Eechaute (149) and Dennis et al (60) had used a Kappa test, they may have produced different outcomes. Secondly, there were only two assessors in both the previous studies, while the current study included eighteen assessors. It

would be reasonable to assume that the inclusion of more assessors might have reduced the level of reliability when compared to that in the previous studies. This is based on the prediction that increasing the assessor numbers would lead to a reduced reliability through an increase in response variability (151).

Three studies have investigated the inter-rater reliability of heel raise and comparable tests. Both the current and previous studies produced similar results, and concluded that the inter-rater reliability of the tests were at a clinically acceptable level.

15.6.4 Intra-rater reliability between assessor groups

Intra-rater reliability of each test was assessed from categorical results obtained from all eighteen assessors who each completed the assessment three times. The three groups of assessors were- six musculoskeletal physiotherapists, six sports medics, and six New Zealand army medics. The between assessor group comparison of means produced p-values that suggested there were no significant differences in the intra-rater reliability results (

Table 15-4 15-4). Thus, these data were pooled and are presented together.

15.6.5 Intra-rater reliability

As with the inter-rater reliability results, all three tests reached the minimum acceptable level of agreement for intra-rater reliability; reasons for this are as described for inter-rater reliability above. In addition, with the exception of Test 3, the reported high frequency of clinical use of these and similar tests might also have possibly influenced the agreement levels observed, based upon feedback gained from several assessors from the current study and previous research (160).

The intra-rater reliability for categorical results was Kappa=0.920 (95% CI 0.882 to 0.959) for Test 1, Kappa=0.890 (95% CI 0.839 to 0.941) for Test 2, and Kappa=0.940 (95% CI 0.844 to 1.000) for Test 3. Taking the clinically acceptable minimum value for agreement

as being Kappa greater than 0.4, with Kappa greater than 0.8 being almost perfect (216), the overall level of agreement for all these tests are encouraging, and a narrow confidence interval suggests that this finding was very consistent across the assessors. A more in depth discussion of these results follow.

15.6.5.1.1 Test 1 (Hindfoot angle)

Pownall et al (222), a study identified previously, investigated the intra-rater reliability of a visual estimate of the hindfoot angle (in degrees). They found an ICC>0.92 for two assessors and two repeated trials, concluding that the intra-rater reliability was at an acceptable level. The results from the Pownall et al (222) study are therefore consistent with the results from the current study, both of which investigated the reliability of a visual estimation of hindfoot angle.

A study by Haight (68), which has also been outlined previously, also investigated the intra-rater reliability of a visually based hindfoot angle assessment. The current and previous studies are similar in that they included both a physiotherapist and a sports medic practitioner. Intra-rater reliability of these measurements was reported by Haight (68) as 'excellent', with ICC of 0.88 to 0.98. This outcome is similar to the agreement reported in the current study, reporting a Kappa value that was 'almost perfect' (151). However, it is important to note that there were a number of differences between the two studies (as discussed previously), including the stance position used, outcome reporting, number of included assessors, and most importantly, that the Haight (68) study was not adequately powered.

Another way in which the studies differ, which may have influenced the results, is the number of repeated trials. The previous study included only one repeated trial, therefore two trials in total, performed by two assessors. The current study included two repeated trials, three in total, and a total of eighteen assessors. It is therefore reasonable to argue that the

current study utilised a more rigorous investigation of the intra-rater reliability of visual assessments.

Two studies to date have investigated the intra-rater reliability of hindfoot angle, measured visually, both of which to an acceptable level of methodological rigor (based on the inclusion criteria presented in Chapter 2, and that the current study is powered to 80%). Both the current and previous study, despite their differences, found that intra-rater reliability was 'almost perfect' and therefore its continued use within the clinical setting is justified, subject to the limitations in both studies, as discussed later in the chapter.

15.6.5.1.2 Test 2 Single legged stance

Haupstein and Goldie (58) investigated the intra-rater reliability of a single leg stance performance. Their methodology was similar to the current study: videotaped performances (N=20) of single limb stance were graded by clinically experienced physiotherapists (N=14), according to an 11-point rating scale. The test-retest reliability was found to be ICC=0.88, and was therefore at an acceptable level.

Finnoff et al. (145) also investigated the intra-rater reliability of single limb stance, within a balance error scoring system (BESS) protocol. For their study, participants were required to stand on one leg on a firm surface for twenty seconds, and given demerit points based on an observable, objective list of errors. Thirty participants and three observers were included in the study and each performed the test twice. The ICC for single leg stance assessment was 0.88 (95% CI 0.76 to 0.94), which is similar to the intra-rater reliability agreement reported in the current study (Kappa =0.890; 95% CI 0.839 to 0.941). There are methodological similarities between the current study and that performed by Finoff et al. (145) which may explain the similar results. Firstly, both involved a visually assessed objective scoring protocol, which required participants to stand on their leg of choice for a twenty to thirty second time frame. Secondly, both the current and previous study performed

a power calculation, which as discussed would limit the chance of a type two error: a real association being missed.

There were, however, some key differences between the two studies, the implications of which have been discussed previously. These differences include the number of assessors, repeated trials and the use of an ICC test to determine reliability. The current study found that Test 2 had a clinically acceptable, 'almost perfect', level of intra-rater agreement, which is consistent with the findings of previous studies (145). Thus, three independent studies, two of which were powered to 80%, have investigated the intra-rater reliability of single leg stance based on visual measurements. All have found that intra-rater reliability is high, and above the minimum clinically acceptable level.

15.6.5.1.3 Test 3 (Heel raise)

The study by Dennis et al (60), as discussed previously, also investigated the intra-rater reliability of a heel raise test. They found that the intra-rater reliability for ten participants, two assessors and two trials was ICC>0.8. This was above the authors' minimum acceptable level. The study by Eechaute (149), also discussed previously, investigated the intra-rater reliability of the multi-hop test and reported ICC values of over 0.9 for two observers over two trials with fifty eight participants. Thus, both the current and previous studies support the intra-rater reliability of tests of the heel raise, and at a similar level.

To summarize, the current study investigated intra-rater reliability of a visually based assessment of heel raise according to the protocol used within FootFAST. This study found that Test 3 had an almost perfect level of intra-rater agreement; Dennis et al (60) investigated the intra-rater reliability of a calf raise test and produced similar results. Further to this, a comparable previous study (149) found that intra-rater reliability of a similar test was high and thus justifies its continued use.

15.6.6 Criterion based validity

Criterion based validity was assessed using the average assessor score (eighteen assessors) for each of the eighteen participants, and for all three tests. This was compared with the relevant 'gold standard' assessment (in accordance with the definition of criterion based validity adopted from Dean-Brown (141)). For Test 1 the assessor scores were compared with hindfoot angular measurements made using Silicon Coach™. Test 2 and 3 were compared to COP movements tracked using the RS Footscan™ plate and Gait™ software during the same performances utilised for the assessor diagnosis.

Test 1 and Test 2 had criterion based validity that reached a clinically acceptable level; in contrast, results for Test 3 were marginal. There are potentially several reasons for this finding, including the lack of similar tests encountered in routine clinical practise. Thirteen of the assessors (13/18) had not previously encountered the Test 3 protocol or scoring system presented here; the remaining five indicated that although they were somewhat familiar with the protocol, they had never used the allocated scoring system. In order to reduce the effect of such unfamiliarity in future research and clinical practice, it is recommended that assessors undergo a training session prior to using the protocol to ensure that they become comfortable with, and confident in the use of, the assessment (146). This would have been incorporated as part of the current study, had the issue been identified prior.

Secondly, there was difficulty faced in deciding upon an appropriate software based comparison method. No methods for validation comparisons between a heel raise test and a 'gold standard' had been previously reported. Thus, the software based analysis in the present study was the first of its kind to be used as a validation comparison. It may be the case that the choice of comparison was not optimal, but it was presumed to be the closest approximation to that which is currently available. No alternative, validated tool had previously been suggested or reported upon: the issue is then whether there is a lack of

validity in Test 3, or whether it is simply that the comparison test was inappropriate. This issue is discussed further in the Conclusions chapter (Chapter 8).

Finally, some smaller perturbations of the foot, seen by the assessors and used within the scoring procedure, may not have been of sufficient magnitude to cause deviances in the COP path viewed within the RS Footscan Gait^{TM} software system. The RS Footscan has not previously undergone such trialling and thus this is speculative at this point. However, it was noted by the author that when viewing the video footage of the participants, notably more perturbations were evident than could be viewed during COP tracking. Thus, it would be possible that the assessors (correctly) allocated a worse score than that gained during the software based analysis. This suggestion could be tested by simulating a heel raise performance, and time-marking instances were small amounts of visual perturbations are apparent, then viewing the same performance within the Gait^{TM} software and comparing.

15.6.6.1.1 Test 1 (Hindfoot angle)

The criterion based validity assessment of Test 1 produced a Kappa score of 0.448 (95% CI 0.375 to 0.521), and thus is above the clinically acceptable minimum of 0.4 (216). A relatively narrow confidence interval gives confidence in the results.

Haight (68) also investigated the criterion based validity of the visual assessment of hindfoot angle compared with a goniometer measurement of hindfoot angle. The agreement level reported was ICC 0.64 to 0.95. This result is consistent with the results gained in the current study, as both outcomes suggest that the visual assessment of hindfoot angle has acceptable validity when compared with goniometry and CAD analysis. The similarities and differences between the two studies have been discussed previously.

An additional difference relates to the comparison measurements. The goniometry measurements obtained in the previous study were calculated by the same assessors who performed the visual assessments. This is arguably not appropriate given the potential for bias

in this situation where the assessors would potentially be able to identify the participant and their own previous assessment. There were a number of steps taken to ensure this potential for bias was minimised in the current study. Firstly, the assessors were provided with only pictures or videos of the participants, which included only from the knee distally in the field of view. This minimises the chance of an assessor remembering the individual participant. Secondly, when the repeated tests were performed, the order of the pictures and footage was randomised using a sequence generated by Microsoft Excel™. Finally, the 'gold standard' comparison measurements obtained were interpreted by the researcher (i.e. not an assessor), and this was performed prior to receiving the results from the assessors.

Beyond this, use of a goniometer as a gold standard comparison is also not considered to be the most appropriate choice. When investigating criterion based validity, a visually based assessment is recommended to be compared to an objective assessment (141). The nature of the goniometer means that there is a substantial amount of subjectivity, and thus potential for bias, involved in the measurement (224): for example, how the goniometer is placed with respect to anatomical reference points. Both this point and the previous indicates that there is an increased risk of bias in the Haight (68) study, which potentially affects the results obtained.

Frigg (150) also investigated the criterion based validity of visual assessments of hindfoot angle. Three different assessors completed a right and left foot assessment on ninety eight participants, and results were compared to radiographic assessments. That study is not entirely comparable to the current study as hindfoot angle was assessed during walking from information on kinetics generated by a Pedobarograph, and compared with radiography. However, given the lack of comparable literature available, it is relevant to note that they found a 48% agreement level between visual and radiography based assessments, which is similar to the level of reliability found in the current study.

A significant limitation of the Frigg (150) study was their comparison of a dynamic assessment with static radiography; the justification for this was not made entirely clear within the study. In the current study both measurements were made statically from the same picture. Thus, it can reasonably be presumed that the comparison made during the current study is more appropriate in a comparison of validity between two types of assessment.

Three studies to date have investigated the criterion based validity of a visual measurement of hindfoot angle. All of these were to an acceptable level of methodological rigor. Both the current and previous studies, despite their differences, found that criterion based validity was of an 'acceptable' standard and therefore the continued use of the test within the clinical setting is supported.

15.6.6.1.2 Test 2 (Single leg stance) and Test 3 (Heel Raise)

The criterion based validity for Tests 2 and 3 were Kappa=0.830 (95% CI 0.750 to 0.911) and Kappa=0.390 (95% CI 0.334 to 0.447) respectively. Thus, the overall level of agreement for Test 2 was 'almost perfect', and for Test 3 this approached, but did not meet, an acceptable level (216). The narrow widths of the confidence intervals for both tests indicate that this finding was consistent across comparisons. This may suggest that for Test 3, the overall Kappa of 0.390 is a reasonably accurate measure of the tests criterion based validity when compared to Gait^{TM} software.

The study by Haupstein and Goldie (58) discussed previously also investigated criterion based validity of a visual assessment of single limb stance, and compared it with results from a force platform. They found criterion based validity to be high, with Pearson's r values of 0.84 to 0.83. The comparison in the previous study was highly comparable to the comparison made in the current study. Both studies identified that criterion based validity was above a clinically acceptable minimum.

Unfortunately, no further criterion based validity studies of single leg stance were identified. No previous studies of a visual assessment of heel raise were found within the literature. However, a previous study, which was identified as being of high quality (Chapter 2), is broadly comparable with the results of both Test 2 and Test 3. Perret et al (54) investigated the criterion based validity of measuring pelvic mobility with the Rachimetre compared with radiography assessment. Although the movements and comparisons are different, both studies assessed comparisons between a subjectively-graded, functional movement (non-static) and an objective measure.

Perret et al (54) identified an ICC of 0.89. This is a similar result to Test 2, and a more positive result than the outcome of Test 3. Both the previous and current study have identified that certain dynamic, functional movement tests can have an acceptable level of criterion based validity. This is an encouraging finding, as it may have been assumed that the more complicated nature of the dynamic functional assessment would potentially compromise criterion based validity. This would have been a concern, considering the number of such tests available to, and used by, practitioners (23).

As there were no previous criterion based validity studies of single leg stance, or of a visual assessment of heel raise, found within the literature there are no means through which to directly compare the results obtained. Despite this, the findings of this study, which is the first of its kind, are encouraging and helpful in the guidance of future research.

15.6.7 Summary of psychometric findings

Table 15-7 shows a collation of both current and previous findings. The ‘interpretation’ column of the results shows positive results: for all psychometric properties investigated the results are ‘acceptable’ or better, with the exception of the criterion based validity investigation of the heel raise test performed in the current study. These findings indicate that the choice of the three tests which comprise FootFAST were appropriate, as in a general

sense, the findings of investigations into their psychometric properties consistently produces clinically acceptable results. The clinical and practical relevance of these findings are discussed below.

Table 15-7: Comparison of current results with the results of previous research

Inter-rater reliability				
Study	**Test**	**#Assessors/ #Participants**	**Outcome**	**Interpretation**
Current	Hindfoot angle	18/18	Kappa=0.450	Acceptable
Haight **(68)**	Hindfoot angle	2/18	ICC=0.56-0.65	Acceptable
Current	Single leg stance	18/18	Kappa=0.450	Acceptable
Trojian **(153)**	Single leg stance	2/40	Kappa=0.898	Almost Perfect
Finnoff **(145)**	Single leg stance	3/30	ICC=0.83	Acceptable
Current	Heel raise	18/18	Kappa=0.450	Acceptable
Dennis **(60)**	Heel raise	10/2	ICC>0.8	Acceptable
Eechaute **(149)**	Multi-hop test	2/59	ICC=0.9	Acceptable
Intra-rater reliability				
Study	**Test**	**#Assessors/ #Participants/ #Trials**	**Outcome**	**Interpretation**
Current	Hindfoot angle	18/18/3	Kappa=0.92	Almost Perfect
Haight **(68)**	Hindfoot angle	2/18/2	ICC=0.88-0.98	Acceptable
Current	Single leg stance	18/18/3	Kappa=0.89	Almost Perfect
Finnoff **(145)**	Single leg stance	3/30/2	ICC=0.88	Acceptable
Current	Heel raise	18/18/3	Kappa=0.94	Almost Perfect
Dennis **(60)**	Heel raise	10/2/2	ICC>0.8	Acceptable
Eechaute **(149)**	Multi-hop test	2/59/2	ICC=0.90	Acceptable
Criterion based validity				
Study	**Test vs Comparison**	**#Assessors/ #Participants**	**Outcome**	**Interpretation**
Current	Hindfoot angle vs Silicon Coach	18/18	Kappa=0.45	Acceptable
Haight **(68)**	Hindfoot angle vs Goniometry	2/18	ICC=0.64-0.95	Acceptable
Frigg **(150)**	Hindfoot angle vs Radiography	3/98	48% agreement	Acceptable
Current	Single leg stance vs RsScan Gait	18/18	Kappa=0.83	Almost Perfect
Current	Heel raise vs RsScan Gait	18/18	Kappa=0.39	Poor
Perret **(54)**	Pelvic mobility vs Radiography	2/10	ICC=0.89	Acceptable

15.7 Clinical relevance

The justification of functionally-based assessments of injury risk has been somewhat controversial, given that it is such an under-researched area. This point is highlighted in the

discussion within Chapter 3. The results of the current study provide evidence for the continued use of Test 1 and Test 2, and the revision of Test 3 (as will be discussed in the section below titled '*Future research directions*'). Future research should therefore be directed at identifying which other functionally-based assessments of injury risk are justified in their current form, which need revision, and which are not justified for continued use. All three of these scenarios are equally important and relevant to the clinical practitioner. Furthermore, the usefulness of such tests across distinct populations, such as athletic groups, the general public and the military must also be investigated. The current study has focussed on the military as a target population, therefore continued research utilising FootFAST with other populations is needed.

The knowledge that Test 3 does not meet the minimum standard for criterion based validity highlights the need for further development of the test. There are a number of potential options to explore in this regard: restructuring of the measurement process used within the test, considering why the test is used and how it is scored, investigating the criterion the test is scored against, or perhaps changing the comparison for the criterion based validity assessment. One option for the latter is a time-to-boundary assessment, which is used to quantify the spatiotemporal characteristics of postural control (59). Time-to-boundary assessments have been proven to be valid, and detect postural unsteadiness that traditional measures cannot (59, 225). Further to this, the intra-session reliability of time-to-boundary measurements have been shown to be comparable to COP measurements used in the current study (59). Given the nature of the instructions provided to the assessors, which focus on time-to-boundary aspects of balance, such as losing balance off the inside or outside edge of the foot, it is possible that the time-to-boundary assessment may be a more appropriate comparison to use. Although it is not the purpose of the current thesis to assess this further, it is suggested that an alternative test, such as the single-legged hop for example, may perform

better in a test of criterion based validity. This suggestion will be considered further, later in the Chapter. However, for now, the current findings highlight the importance of continued research in this area.

In conclusion, based on the findings from the current study, with respect to psychometric properties, Test 1 and Test 2 are suitable for continued use within the clinical environment in their current form.

15.7.1 For practitioners

Results from this study have important implications for clinical practitioners. As has been discussed in the literature review chapter (Chapter 3), it is hard to justify the continued use of a test when the psychometric properties are unknown. Before now, none of the three tests within FootFAST had been investigated for their clinometric properties. The current research study has high ecological validity as the tests included in the study were performed to reflect, as far as possible, how they would be performed in the clinical setting. Of course, it is lacking in that a full clinical history of the patient was not provided to assessors, and in a clinical setting there would be no blinding. These points aside, all three tests which comprise FootFAST demonstrated acceptable levels of both inter and intra-rater reliability; thus, an important step in justifying their use in clinical practise has been made. Two out of three tests had acceptable validity, the exception being Test 3.

There are several practically relevant outcomes arising from the findings. Firstly, the New Zealand army are currently investigating potential overuse injury risk assessment tools for use among new recruits. FootFAST was trialled within the New Zealand army in 2009, and the findings suggested that FootFAST had substantial ability to predict overuse injuries of the lower limb among recruits (1). Beyond this, the use of NZ army medics as one of the assessor groups in the current study has provided evidence to the suitability of its use within this environment. The current study found that NZ army medics performed the risk

assessment protocol to the same standard as both physiotherapists and sports medicine practitioners. Thus, if the army were to use FootFAST as their risk assessment protocol for new recruits, they would not need to use external or specialist-trained examiners.

This knowledge is important for sporting groups too: while a physiotherapist may previously have been needed to perform injury risk assessment protocols (typically utilised pre-season), this would not be the case for FootFAST screening. Where the cost and practicality of such screening by a physiotherapist could potentially have been a barrier, if a sports medic is available, the current results demonstrate that they can also perform such assessments to a clinically acceptable level. This potentially provides more opportunities for sporting and athletic teams in terms of risk assessment and the identification of injury predictors.

15.8 Strengths and limitations

15.8.1 Strengths

The main strength of the current study is in its novelty and rigorous methodology adopted. No previous studies have investigated in completeness the inter-rater reliability, intra-rater reliability, and criterion based validity of any of the three tests included in the current study. Indeed, the current study has produced new research and information in an area which is lacking. Secondly, the use of a feasibility study (Chapter 6) ensured that the main study for this thesis had sufficient power to assess the inter-rater and intra-rater reliability and criterion based validity. This is important, as inadequately powered studies have been prevalent in the field: few studies are powered to investigate the psychometric properties of clinical tests (accepted as 80% or more).

15.8.2 Limitations

The limitations of the current study, given that the main purpose was to investigate the psychometric properties of each of the three tests within FootFAST, are discussed below.

A minor limitation of the current study was that the participant group only included males from a relatively narrow age range. In some ways this may be considered a strength, as it reduced the possibility of sex as a confounder, and because the group which they are sampled from is made up of over 80% males (Lt Col Dunn, Personal Communication). However, it does reduce the generalizability of the results. No previous studies have investigated the potential differences between males and females in such performances, with the exception of Test 2. Schneiders et al (160) found that there were no differences between the performances of males and females in a single leg stance test. Thus, the extent to which including only males may be limiting the generalizability of Tests 1 and 3 remains speculation at this point. However, based on the findings of Schneiders et al (160), it is likely to be unproblematic.

Limitations which may have contributed to Test 3 failing to produce clinically acceptable levels of criterion based validity include the study's methodology. Assessors were provided with images or videos of participants which contained only the lower part of the shin and foot from the rear. Although this procedure increased the rigor of the methodology (i.e. all assessors had access to precisely the same image), it reduced the ecological validity of the assessment. In a clinical setting, the patient would be in the same room as the assessor whilst the performance was judged. In this situation, it is possible that an assessor would observe the patient from different angles, and perhaps even use observations of their patient in motion before and after the performance to assign a grade. In addition, in the clinical setting, the assessor would have access to subjective information from their patient: lack of

such information in the current study may have also affected the psychometric performances of all three tests within FootFAST.

The category choices presented to the assessors as classification options may also have been problematic. The grading system used within FootFAST was developed by its creator (Dr. Charles Baycroft) and therefore is unique. None of the assessors had used these classification systems before. While similar classification systems of various types are commonly used in clinical practice, the author could not find any previous test using the same classification systems, making such systems unique and specific to FootFAST. Familiarity with a test is likely to increase the consistency of its performance (146), therefore, the lack of familiarisation with FootFAST may have led to different interpretations of the categories presented, and individual assessors applying their own criteria as part of the grading. It could be argued this is supported by the difference in Kappa scores between inter-rater reliability and intra-rater reliability (for example, in the assessment of Test 1 these were K=0.450 and K=0.920 for inter and intra-rater reliability). Although this was not directly investigated here, questions pertaining to this could have been included as part of the feedback obtained from the assessors.

Finally, the gold standard comparison chosen for each of the tests may not have been optimal: there are currently no universally accepted ‘gold standard’ assessments for the tests investigated here. With reference to Test 1, it is commonly accepted in the literature that a radiography scan or x-ray is the gold standard comparison for the assessments of hindfoot angle (150). However, the major problem with radiographic assessment is the attendant exposure to radiation, and its associated risks (226). Use of radiographs are also impractical for routine use: for the current study, this would have required additional funding and presented significant logistical and practical issues as there was no radiology department in close proximity to Burnham Military Camp. Imaging using ultrasound scanning would not

necessarily provide a better alternative. Finally, the non-weight bearing position of the hindfoot as viewed in a scan may be quite different to the position of the hindfoot when weight bearing, and thus reduce the validity of a scan as a comparison to visual assessments made when standing. This suggestion is supported by findings of previous research: Nosewicz et al (227) found that radiology scans had only poor to moderate validity in the assessment of the talus in sagittal and horizontal plane measurements.

Based on these factors, for the purposes of the current study it was considered that comparison with CAD measurements was an appropriate alternative. 2D CAD measurements, as utilised here, have been shown to have excellent reliability (ICC>0.91) when reflective markers are used (38). Reflective markers were included in the current test when CAD assessments were made. Further, these 2D CAD measurements were found to be highly correlated (Pearsons r = 0.95) with manual, goniometric measurements (38). The extent to which underlying measurement variability, despite reliability being at an excellent level, may have affected the current results is unknown.

For Tests 2 and 3, there is no universally accepted gold standard comparisons specifically identified in the literature. In the case of Test 2, various protocols have been used previously (for examples refer to Meijne et al (137) or Trojian (146)). The comparison used in the current study was chosen due to pragmatic considerations, and the informed opinions of the research team. However, the use of COP tracking as a comparison could be disputed, as there is not a universally accepted measure to interpret the outputs gained. For the purposes of this study, it was considered that the use of RS Footscan™ pressure plate measurements was an acceptable instrument to be used in lieu of a gold standard comparison.

Another method which could have been employed in the current study is the transformation of the RS Footscan™ outputs, in order to smooth the data. Measuring the acceleration of the COP, then smoothing the data using a root mean squared transformation

demonstrates one possible method. This has been used successfully in previous studies to measure postural control (228), and is said to make the interpretation of the outputs more simplistic, as they could be observed and interpreted without the need for further analysis. It is unknown whether the use of such a transformation may have produced better criterion based validity scores than those observed in the current study.

Despite the issues identified, comparisons of Tests 2 and 3 with RS Footscan™ plate system were considered by the author (who undertook the comparison) to be unproblematic with respect to practicality and feasibility. However, there has been no research to date which investigates the repeatability of RS Footscan™ pressure plate measurements, and thus the extent to which any such variability may have affected the current results is unknown.

15.9 Future research directions

15.9.1 Revision of Test 3

It was found that one of the three tests within FootFAST (heel raise) did not meet the minimum acceptable level of criterion based validity. Future research should be directed at either modifying the protocol for Test 3 or searching for an alternate means of assessing function in this way (see below). The results of this study have indicated that Test 3, as measured using the protocol described in the methodology section, is not justified for continued use in isolation.

Apart from a lack of validity, one of the potential reasons that Test 3 scored poorly in the criterion based validity assessment is in part due to the choice of comparison test which was potentially not optimal for the purposes of this study. As there had been no previous research which had investigated the heel raise test in such a way, the choice of test was based on the informed opinions of the author and supervisory team. Thus, it is possible that the comparison, given its novelty, was not optimal. The test produced acceptable levels of both inter and intra-rater reliability, despite the assessors lack of familiarisation with the test,

which supports the continued use of the current procedure and scoring system. Further to this, a recent study by Silbernagel et al (229) found that the heel raise test for height had a high level of construct validity when used in isolation (differences in functional performance between injured and non-injured ankles were statistically significant ($p<0.02$)). Thus, criterion based validity appears to be the only aspect of psychometric properties where Test 3 is lacking. The discussion then leads to what other 'gold standard' tests could be used as comparison?

An alternative comparison, which was also considered at the beginning of the study, is use of a force plate instead of the pressure plate system. The force plate may provide more accurate information regarding the direction of forces applied by the foot during the heel raise, due to the increased sampling speed of the plate (Kistler™ user manual 9281A). Further to this, a recent study by Veilleux and Rauch (230) found that the co-efficient of variation (measurement of reliability) when using such a force plate to assess a heel raise was <7.5% (twenty-eight participants, two sessions), which is described as being 'highly reproducible'. This adds further support to its use as an alternative to the pressure plate.

In contrast, the pressure plate system provides information concerning the centre of pressure movements (RS Footscan™ manual) rather than force vectors. Direction and size of perturbations in movement, which is essentially what the assessors were grading performance on, can be interpreted through force vectors. This can be derived from the COP information provided by the RS Scan, but with moderate reliability (231). The reasons for not utilising the pressure plate in the current study were based on availability and practicality. As the participant sample was recruited and measured on-site at Burnham Military Camp, the portable nature of the RS Scan was more practical. In addition, the force plate available to the researcher (in common with most systems) did not have the sensitivity (in terms of acuteness of measurements) of the RS Scan, which suggested the RS Scan could be more appropriate

for the current study. Based on the results of the current study, future research is needed to investigate the comparison of visually based assessments and force plate measurements, and *inter alia* to assess its potential as a more appropriate comparison for Test 3.

As indicated earlier in the Chapter, another option is to potentially replace Test 3 (currently a heel raise test) with a more appropriate alternative. The incorporation of the heel raise test was to ensure that all components of the weight bearing gait cycle are investigated within the FootFAST protocol. The specific aim of including the heel raise test was to provide a correlate with the toe-off phase of the gait cycle. In replacing the test, it would be important to ensure that the alternate test provides an indication of function during the same phase. As such, the single leg hop test has potential to be utilised within FootFAST.

The single leg hop test involves the same action as the heel raise, in that it involves a unipedal plantarflexion, and requires the utilisation of various postural strategies in order to retain balance (149). The main difference is that the increased speed of plantarflexion involved in the single leg hop requires a greater degree of postural control given the dynamic nature of the test. It can be argued that this test is perhaps a more ecologically valid assessment of the toe-off phase, and thus, is more natural to a participant, because of the rapid movement and the forwards translation. Although there have been no investigations of the psychometric properties of a 'single' one-leg hop test presented in the literature, a repeated one-leg hop test has been studied by one group (149). The outcome of this research supports the suggestion of including a single one-leg hop within FootFAST:

As previously discussed (Chapter 2), Eechaute (149) investigated the inter and intra-rater reliability of the multi-hop test: results indicated that the test may perform comparatively better in both inter and intra-rater reliability (ICC greater than 0.90 for both) than Test 3. Unfortunately, there were no comparisons found which investigated criterion based validity of the single one-leg hop test. Future research should therefore investigate the

criterion based validity of the single one-leg hop test, and confirm the inter and intra-rater reliability of the test. Interestingly, Veilleux and Rauch (230) also included the single one-leg hop test in their investigation of reliability of force plate measures. They found that the co-efficient of variation for this test (twenty-eight participants, two sessions) was also <7.5%. Thus, the option of using the force plate as a criterion based validity comparison for the single one-leg hop test is also an option based on these findings. Following this work, the potential for the test to be included in FootFAST can be considered.

15.9.2 Participants

There are still many areas to address in the future research of FootFAST. These include the patient populations: FootFAST research to date has only included New Zealand army soldiers. Therefore, other populations including sporting groups and the general population need to be included if the screening test is to be considered for wider populations. For example, it has been shown in previous studies that as the age of the population group increases, their balance ability reduces (190, 228). Given that two of the three tests within FootFAST are largely affected by the participants balance ability (single leg stance and heel raise), it is speculated that older populations would perform poorer than the population utilised in the current study. This notwithstanding, a poorer performance in FootFAST may, or may not, produce better reliability scores among assessors. No research was found which had directly investigated this, and thus future research is needed in this area.

Further, a feasibility study of the predictive ability of FootFAST has only been investigated in a single study to date, which again included a participant sample from only New Zealand army soldiers (1). Although the results of this study were positive (statistically significant reductions in eight out of ten overuse injury rates), further research is needed to confirm these successful results, and to explore the effectiveness of FootFAST as an injury risk assessment within other populations.

15.9.3 Assessor training

As discussed in Chapter 6, none of the assessors included in the current study had been trained in using the FootFAST protocol, or had used FootFAST previously. This ensured that all the assessors had a standardised amount of experience in using the protocol. Although this improved the methodological quality of the current study, the results gained may not be replicated in an assessor group who is experienced in using FootFAST. Haight (68) found that novice assessors had poorer reliability than experienced assessors, therefore, it could be speculated that an experienced assessor who is using a protocol for the first time may have poorer reliability than an experienced assessor who is familiar with a protocol. As this has not been previously investigated, future research is needed in order to confirm this suggestion.

15.9.4 A critique of FootFAST

An important aspect in the development of tests such as FootFAST is a critique. There has currently been no such process undertaken, which represents an important area for future research. Despite that the three tests scored well in the assessments of inter and intra-rater reliability, and criterion based validity (with the exception of Test 3), there may be tests which are more appropriate to include as part of the tool. Evidence to support this suggestion was identified in Chapter 3, which found several functional tests with higher scores of reliability and/or validity (as discussed). Further to the tests performance in these areas, there are other aspects of a clinical test which also need to be considered, which were not directly assessed in the current study. For example, if there are tests available which are more sensitive to injury risk, or more sensitive to changes in injury risk over time, then these options need to be explored. As an investigation of sensitivity of the three tests within FootFAST was not included in the current study, future research is needed in order to make this assessment.

15.9.5 Implications and recommendations for clinical practise

The final point for discussion is how to provide recommendations for groups such as sports teams, or the New Zealand army, for future application. The utilisation of FootFAST as a tool for assessing injury risk is supported by current research. Results of the current study also indicate that practitioners who are easily accessible to both sporting teams and the New Zealand army can effectively perform the investigation without the need for additional training. This has clear implications in terms of access and affordability of injury screening for these groups.

This notwithstanding, it is recommended that when utilising FootFAST before changes are made to Test 3, some caution is exercised in interpreting the findings. It should be acknowledged that only assessors with a minimum of three years of clinical experience were included in the current study, and therefore similar reliability may not be achieved when novice practitioners perform these tests (as was the case in previous investigations (68)).

Secondly, although the current research has a relatively high level of ecological validity, a limitation was that the assessors were only provided with digital images, and did not complete a real-time assessment. While a person-to-person assessment might produce better results, it is still unknown whether this would be the case. Previous research has indicated that the agreement between video and real-time assessments performed by physiotherapists is highly variable (232). It may be that because the assessors in the trial were able to replay the images as many times as they wish, and utilise ‘pause’ and ‘slow time’ features, that the performances were easier to grade. Future research should be directed towards investigating these assumptions when FootFAST is performed under clinical conditions (i.e. ecologically valid settings).

15.10 In the next Chapter

The current Chapter presented the main study of the thesis, which investigated the psychometric properties of the three tests within FootFAST, incorporating the knowledge gained during the feasibility study presented in Chapter 5. The next Chapter concludes the thesis by briefly summarising the findings.

16 Chapter 8. Thesis Summary

The purpose of the current thesis was to determine whether the FootFAST clinical assessment protocol was justified for continued use in a clinical environment, based on its scores of intra-rater and inter-rater reliability, and criterion based validity. FootFAST is an overuse injury risk assessment, which focuses on the bipedal functionality of the lower limb. A previous observational cohort study by the author identified that FootFAST could be used to identify NZ army recruits at risk of overuse injury, and inform appropriate orthoses management. However, the psychometric properties of FootFAST were unknown, and the effectiveness of the intervention utilised (orthoses) had never been assessed. Thus, the purpose of the current thesis was to address these issues. In order to achieve this, the following five research questions were asked:

1. What is the current evidence of clinical effectiveness of orthoses in reducing target injuries in the population of interest: the military?
2. What is the current evidence for psychometric properties of clinical assessment tools for injury risk identification?
3. What is the inter-rater reliability of the three tests within the FootFAST screening protocol?
4. What is the intra-rater reliability of the three tests within the FootFAST screening protocol?
5. What is the criterion based validity of the three tests within the FootFAST screening protocol?

16.1 Research Question 1

In order to answer the first research question, a systematic review including meta-analysis was undertaken (Chapter 4). There was no evidence found which either supported or refuted

the use of orthoses as an intervention tool, and therefore it is suggested that future studies investigate alternate intervention options. Although there was no evidence of harm from the use of orthoses interventions to reduce injury risk, the lack of evidence of their effectiveness questions their current use, including their purpose, effectiveness and – in turn – cost-effectiveness.

An issue highlighted during the systematic review was that many of the trials retrieved were not of sufficient quality to allow firm conclusions to be made, despite the fact that approximately thirty relevant studies were identified. This issue was also identified in the narrative review (Chapter 3), and is of concern, as a lack of quality studies is in many ways as problematic as a lack of studies. Future researchers should focus on optimising their research methods in order to ensure the robustness and validity of their findings. The main methodological shortcomings included the removal of non-compliant participants from the results analyses, which creates an intention-to-treat bias. Also, there was a lack of detail in the methodologies which did not allow for a complete assessment of potential bias in the studies reviewed.

The systematic review focused investigation on military populations. This review added to the current field of knowledge, no such review had been undertaken previously. It is considered important that future reviews of this kind also focus on specific participant groups. As outlined in Chapter 4, it is not entirely appropriate or relevant to pool results from studies which include varying groups of participants. Across different subgroups, exposures to risk and physical loadings will be different, and thus, exposure differences are the primary reason behind the suggestion.

Based on the systematic review, it was found that there was insufficient evidence on injury prevention capabilities of the orthoses, and an absence of cost-effectiveness analyses within the included studies. Thus, there was no compelling evidence to suggest that orthoses

use is appropriate as a cost-effective intervention for injury prevention. In addition, there was also no evidence to suggest that orthoses caused harm.

16.2 Research Question 2

The second research question concerned current knowledge regarding the psychometric properties of visual assessments of function. Given the limited numbers of research studies, and the heterogeneity of studies and information, a systematic review and meta-analysis was not possible. Instead, the findings were presented in a narrative review. The most important finding from this investigation was that the popularity of functional assessments to identify injury risk is at odds with current levels of knowledge regarding the reliability and validity of these tools.

The narrative review conducted in Chapter 3 added to the current field of knowledge in several important ways. There had been no previous reviews in this area; thus, the knowledge presented in the findings is novel. There were several key issues raised in the Chapter discussion which should be used to inform future research in the area of functional assessments to identify injury risk:

1. Research into the psychometric properties of visual assessments of function for injury risk is lacking, therefore more studies are needed, not least given the widespread use of visual assessments in clinical practice.
2. There is a lack of systematic investigation of psychometric properties: many research investigations concentrate on only one aspect of reliability and/or validity.
3. The quality and rigor of the investigations is disappointing, e.g. many of the studies which investigated inter-rater agreement included only two assessors.
4. There was no consensus on what should be used as a gold standard comparison in investigations of validity. This makes results from several studies difficult to compare and

therefore comparisons can be confusing to practitioners. Thus, it is important that researchers aim to develop, and agree on, appropriate and universal gold standard comparison measures.

5. There was an absence of evidence identified which validated functional screening in the prediction or prevention of injury.

16.3 Research Questions 3-5

16.3.1 Reliability

Question three focused on the inter-rater reliability of Tests 1-3 (hindfoot angle, single leg stance and heel raise). The agreement level was 'satisfactory', and reached the clinically acceptable minimum, Kappa=0.450, across all three tests. The fourth research question aimed to identify the intra-rater reliability of the three tests within FootFAST: agreement levels for the three tests ranged between Kappa = 0.890 and 0.940, representing 'almost perfect' intra-rater reliability. These results demonstrate the excellent ability of the included assessors in reproducing their patient assessments.

The investigation of inter-rater reliability presented was unique in several ways. Firstly, the current study included three groups of assessors likely to undertake these assessments in practise, whereas previous studies had included only two (sports medics and physiotherapists (67)). This increases the external validity of the present findings. Such inclusion criteria also allowed comparison of intra-rater reliability between three different assessor groups. It was found that there were no significant differences between the three groups. This level of comparison has not previously been presented and provides significant and relevant information: it indicates that practitioners from these three groups can perform the assessments to a comparable level, and thus recruiting a practitioner from any of these groups to perform such a test is justified. In addition, it would allow sports teams and other groups access to such injury risk assessments without the requirement of a physiotherapist, thus potentially providing a more cost-effective and feasible option.

Although inter-rater reliability was above the clinically acceptable minimum across all three tests, it is noted that intra-rater reliability scores were consistently better (almost perfect agreement for all three tests). In order to improve on the inter-rater reliability scores, it is suggested that practitioners who wish to utilise the FootFAST protocol undergo specific training. Although the assessors included in the study were either familiar with, or had used very similar tests to, the three FootFAST tests previously, none had specifically used the FootFAST assessment in entirety or the scoring system. As it has been demonstrated that increased training improves the inter-rater reliability scores of visual assessments of function (68), future research is needed to investigate whether FootFAST-specific training can improve on the inter-rater reliability scores presented in the current investigation.

Intra-rater reliability of the three tests was consistently high. This is of course an encouraging finding, and partially justifies the continued practise of assessment, re-assessment in order to identify improvements (as discussed within Chapter 3). The other aspect of this is the tests' ability and sensitivity in identifying change, a property not investigated here and one that needs to be investigated within future research. The blinding and randomisation employed, helped by the use of video footage instead of real-time assessment, had not previously been attempted. This process minimised recall biases, which would have limited the usefulness of the results. In addition, the number of repeated trials, assessors and participants was greater than previous studies. This further adds to the significance of the current investigation.

Interestingly, on a related topic, familiarity and use of similar tests did not seem to have an impact on intra-rater reliability scores. As indicated in Chapter 6, assessors commented that they had utilised tests similar to Test 1 and Test 2 previously, but not Test 3. Despite this, the three tests all scored highly in intra-rater reliability. This is also encouraging, as it implies that novel assessments with which the practitioner is familiar can be repeated to a similar

standard as those tests with which the practitioner is already trained. Previous tests have also found comparable results in investigations of experienced (i.e. trained) versus novice (i.e. familiar) practitioners (68).

16.3.2 Validity

The final research question concerned the criterion based validity of the FootFAST tests. Using the definition of criterion based validity provided by Dean-Brown (141), the investigation compared the results from the assessors to results gained from an objective 'gold-standard' assessment. Two (Tests 1 and 2) of the three tests reached a clinically acceptable level of agreement, with results of Kappa = 0.448 (Test 1) and Kappa = 0.830 (Test 2). Test 3 achieved an agreement score of Kappa=0.390, and thus did not meet a clinically acceptable level.

The investigation into criterion based validity produced varying results. This observation was also noted in the findings of previous studies; in some cases the validity was as low as 48% agreement, while other studies found no significant differences between subjective and objective scores (68, 150, 153), indicating high agreement levels. This comparison concerns how well the visual test results match results from the criterion 'gold standard' test. However, it is suggested that the criterion based validity score is also influenced by appropriateness of the *choice* of 'gold standard' comparison. If an acceptable minimum level of agreement is not achieved, as is the case within this study for Test 3, it is hard to ascertain whether the visual test is lacking validity, or whether the comparison chosen was not optimal. Test 3 has no gold standard comparison that has been used previously, and the choice was based on the informed opinion of the research team.

This represents a considerable challenge for future research in this area, and it is difficult to address this issue satisfactorily. As was briefly discussed in Chapter 3, it is suggested that future research concentrates on the investigation of construct validity when

investigating psychometric properties of visual assessments of function for injury risk: queries such as whether the test measures what it claims to measure (141). Such research is termed the discriminative or predictive validity of an assessment, and focuses on ascertaining whether the visual assessment of function correlates with risk of injury over a specified period of time to establish whether there was a significant difference in injury occurrence based on the outcome of the visual assessment. Investigations of this nature are much more relevant to the practitioner too, as they are essentially a means of justifying the continued use of the test as an indicator of injury risk.

The current investigation aimed to assess criterion based validity. While only two of the three tests within FootFAST met the minimum acceptable level of agreement, based on the discussion presented, it is not considered to be a convincing reason to discontinue the use of Test 3 in the short term, as the results of the assessment gained from the Test are not used in isolation, but along with the results of the other two tests which comprise FootFAST. Given that the Test scored above an acceptable level in both inter-rater and intra-rater reliability, and construct validity has been supported based on the results from a previous study (1), it is suggested that FootFAST can continue to be used in its current form, with due caution, until further development of Test 3 is possible.

16.3.3 Summary

As indicated above, the psychometric properties of FootFAST were previously unknown. The studies conducted in the current thesis were the first of their kind, and therefore make a novel contribution to the field of injury prevention. In addition, the level of methodological rigor and comprehensiveness of the investigation mean that the results gained are likely to be representative of the populations sampled from, as all attempts were taken to minimise bias, and the study held an acceptable level of power. Based upon these findings, practitioners can

continue to use the tests investigated, in the manner presented, and have confidence in the results obtained.

In saying this, the specific nature of the methodology incorporated must be acknowledged. For example, the participants included in the investigation were injury free for the past three months; hence the applicability of these findings to current or previously injured patients is unknown. Secondly, the inclusion of only musculoskeletal specialists does not allow direct generalizability to other specialist groups of physiotherapists (particularly if we consider the differing results gained in the feasibility study). Finally, as mentioned, there has not yet been an investigation of these tests' sensitivity to change in injury risk status and thus further research is required. Further to this, there is no current evidence supporting the three included tests as being the optimal combination for continued use. As such, FootFAST is still not considered to be clinically relevant. This presents another topic for future research.

The three tests within the FootFAST protocol are designed to be predictive of injury based on a combined result. Of course, it is also possible to use each of the three tests separately, as the current study provides results for the three tests in isolation. The exception to this is, of course, Test 3 which did not achieve the minimum acceptable level of criterion based validity. This point has been discussed in the reliability section above. Clearly more research in this area is needed, as the majority of clinical tests similar to FootFAST have not yet been assessed in this way.

16.4 Final comments

The current thesis contains four novel investigations (narrative review, systematic review, feasibility trial and a repeated measures cohort) in an underdeveloped field of research – overuse injury risk assessment. Based on the findings, practitioners can justify their choice of using any of the three tests investigated.

The current thesis has identified guidance points for future research, and highlighted several aspects in which current research is lacking. Specifically, this was in the areas of psychometric testing and intervention trailing within the field of overuse injury prevention. Future investigations in these areas should consider the points identified here, mainly regarding methodological choices and novelty of investigations, as discussed in detail in the relevant Chapters.

17 References

1. Baxter ML, Baxter DG, Baycroft C. Lower limb injuries in soldiers: Feasibility of reduction through implementation of a novel orthotic screening protocol. Military Medicine. 2011;176(3):291.
2. Baxter M, McBride DI, Milosavljevic S, Hendrick P, editors. Lower limb and lower back overuse injury prevention: Evidence from a specialised population. Southern Physiotherapy Symposium; 2011; Queenstown.
3. Hicks C. Research methods for clinical therapists. 4th ed. China: Elsevier; 2004.
4. Baxter M, McBride DI, Cury-Ribeiro D, Hendrick P, editors. Psychometric properties of lower limb and lower back injury risk assessment: A narrative review. Australian Conference of Science and Medicine in Sport; 2013; Phuket.
5. Baxter ML, Ribeiro DC, Milosavljevic S. Do orthotics work as an injury prevention strategy for the military? Physical Therapy Reviews. 2012;17(4):241-51.
6. Baxter M, McBride DI, Milosavljevic S, Hendrick P, editors. A background and justification of the novel injury risk assessment protocol: FootFAST. Australasian Podiatric Societies Annual Conference; 2011; Sydney.
7. Baxter M, McBride DI, Cury-Ribeiro D, Hendrick P, editors. What are the psychometric properties of the injury risk assessment FootFAST? A preliminary study. Australian Conference of Science and Medicine in Sport; 2012; Sydney.
8. Baxter M, McBride DI, Milosavljevic S, Hendrick P, editors. Are visual assessments made by clinicians reliable and valid? Dunedin School of Medicine Colloquium; 2012; Dunedin.
9. Austin WM. Stress injuries and orthotic support. Roanoke: Foot Levelers, 2003.
10. Fletcher RT. Pulheems: A new system of medical classification. British Medical Journal. 1949;January:6.
11. Bowen JD, Salter G. Gait assessment: Springer; 2010. Chapter 6; p.71-86 p.
12. Bird S, Black N, Newton P. Sports injuries: Casues, diagnosis, treatment and prevention. Canterbury: Epsom and Newport; 1997.
13. Valmassy RL. Clinical biomechanics of the lower extremities. New York: Mosby; 1996.
14. Moore G. Atlas of the musculoskeletal examination. USA: Versa Press; 2003.
15. Barton CJ, Coyle JA, Tinley P. The effect of heel lifts on trunk muscle activation during gait: a study of young healthy females. Journal of Electromyography and Kinesiology. 2009;19(4):598-606.
16. Bradshaw EJ, Keogh JW, Hume PA, Maulder PS, Nortje J, Marnewick M. The effect of biological movement variability on the performance of the golf swing in high-and low-handicapped players. Research Quarterly for Exercise and Sport. 2009;80(2):185-96.
17. Franklyn-Miller A, Boyington W. Can the RsScan orthotic system reduce lower limb injury in an initial military training setting. Melbourne: Centre for Health, Exercise and Sports Medicine, 2001.
18. Hughes M. Performance analysis-a 2004 perspective. International Journal of Performance Analysis in Sport. 2004;4(1):103-9.
19. Richardson S, Cooper A, Alghamdi G, Alghamdi M, Altowaijri A. Assessing knee hyperextension in patients after stroke: comparing clinical observation and Siliconcoach software. International Journal of Therapy and Rehabilitation. 2012;19(3):163-8.
20. Whatman C, Hume P, Hing W. The reliability and validity of physiotherapist visual rating of dynamic pelvis and knee alignment in young athletes. Physical Therapy in Sport. 2012;14(3):168-74.

21. van Mechelen W, Hlobil H, Kemper HC. Incidence, severity, aetiology and prevention of sports injuries. Sports Medicine. 1992;14(2):82-99.
22. Franklyn-Miller A. Foot orthoses in the prevention of injury in initial military training: A randomized controlled trial. The American Journal of Sports Medicine. 2011;39(1):30-7.
23. Buckup K. Clinical tests for the musculoskeletal system. Stuttgart: Thieme; 2008.
24. Gross J, Fetto J, Rosen E. Musculoskeletal Examination. 3 ed. Oxford: Wiley-Blackwell; 2009.
25. Zajac FE. Biomechanics and muscle coordination of human walking. Gait and Posture. 2002;16(3):215-32.
26. Ounpuu S. The biomechanics of walking and running. Clinics in Sports Medicine. 1994;13(4):843.
27. Lovejoy CO. The natural history of human gait and posture: Part 1. Spine and pelvis. Gait and Posture. 2005;21(1):95-112.
28. Whittle MW. Gait analysis: An Introduction. 3 ed. United Kingdom: Butterworth-Heinemann; 2003.
29. Plotnik M. A new measure for quantifying the bilateral coordination of human gait: effects of aging and Parkinson's disease. Experimental Brain Research. 2007;181(4):561-70.
30. Plotnik M. Bilateral coordination of walking and freezing of gait in Parkinson's disease. The European Journal of Neuroscience. 2008;27(8):1999-2006.
31. Murley GS, Menz HB, Landorf KB. Foot posture influences the electromyographic activity of selected lower limb muscles during gait. Journal of Foot and Ankle Research. 2009;2(35):9.
32. Murley GS. The effect of three levels of foot orthotic wedging on the surface electromyographic activity of selected lower limb muscles during gait. Clinical Biomechanics (Bristol). 2006;21(10):1074-80.
33. Hume P. Effectiveness of foot orthoses for treatment and prevention of lower limb Injuries: A review. Sports Medicine (Auckland). 2008;38(9):759-79.
34. Hreljac A. Impact and overuse injuries in runners. Medicine and Science in Sports and Exercise. 2004;36(5):845-9.
35. Davidson PL. Examination of interventions to prevent common lower-limb Injuries in the New Zealand defense force. Military Medicine. 2009;174(11):1196.
36. Braunstein J. Magnetic resonance imaging of hereditary hernias of the peroneus longus muscle. Skeletal Radiology. 1995;24(8):601.
37. Ragland III R, Moukoko D, Ezaki M, Carter PR, Mills J. Forearm compartment syndrome in the newborn: report of 24 cases. The Journal of Hand Surgery. 2005;30(5):997-1003.
38. Miles C, Sparkes V. Intra and inter-rater reliability of SiliconCoach when measuring quadriceps angle. Cardiff: Cardiff University College of Medicine; 2007 [updated 10/04/2013]; Available from: http://www.siliconcoach.com/SupportCentres/Generic/Research_Articles_using_siliconcoach.pdf.
39. Chen W-P, Lin C, Tang T. Dynamic simulation and experimental validation of the plantar foot pressure during heel strike. ASB 29th Annual Meeting; August 5th; Ohio2005.
40. Redmond AC. Development and validation of a novel rating system for scoring standing foot posture: The Foot Posture Index. Clinical Biomechanics (Bristol). 2006;21(1):89-98.
41. Abernathy B, Hanrahan SJ, Kippers V, Mackinnon LT, Pandy MG. The Biophysical Foundations of Human Movement. Australia: Palgrave MacMillan; 2004.

42. Bobick AF. Gait recognition using static, activity-specific parameters. Proceedings of the 2001 Computer Society Conference; Atlanta: Computer Society Press; 2001. p. 423-30.
43. Liang Wang L. Silhouette analysis-based gait recognition for human identification. Transactions on Pattern Analysis and Machine Intelligence. 2003;25(12):1505-18.
44. Schwartz MH, Rozumalski A. A new method for estimating joint parameters from motion data. Journal of biomechanics. 2005;38(1):107-16.
45. Hesar NGZ. A prospective study on gait-related intrinsic risk factors for lower leg overuse injuries. British Journal of Sports Medicine. 2009;43(13):1057-61.
46. Lieberman DE. Foot strike patterns and collision forces in habitually barefoot versus shod runners. Nature (London). 2010;463(7280):531-5.
47. Bartlett R. Sports biomechanics: reducing injury and improving performance: Routledge; 2013.
48. Baumhauer JF, Alosa DM, Renström PA, Trevino S, Beynnon B. A prospective study of ankle injury risk factors. The American Journal of Sports Medicine. 1995;23(5):564-70.
49. Beynnone BD, Renstrom PA, Alosa DM, Baumhauer JF, Vacek PM. Ankle ligament injury risk factors: a prospective study of college athletes. Journal of Orthopaedic Research. 2001;19(2):213-20.
50. Payne CB. Effect of foot orthoses on lower extremity kinetics during running: A systematic literature review. Journal of Foot and Ankle Research. 2008;1(1):13.
51. Krebs DE. Reliability of observational kinematic gait analysis. Physical Therapy. 1985;65(7):1027.
52. Vendrely A, Carter R. The influence of training on the rating of physical therapist student performance in the clinical setting. Journal of Allied Health. 2004;33(1):62-9.
53. Saltzman CL. Measurement of the medial longitudinal arch. Archives of Physical Medicine and Rehabilitation. 1995;76(1):45.
54. Perret C, Poiaudeau S, Fermanian J, Revel M. Pelvic mobility when bending forward in standing position: Validity and reliability of 2 motion analysis devices. Archives of Physical Medicine and Rehabilitation. 2001;82(February):221-7.
55. Gorton GE. Assessment of the kinematic variability among 12 motion analysis laboratories. Gait and Posture. 2009;29(3):398-402.
56. Ageberg E. Balance in single-limb stance in healthy subjects–reliability of testing procedure and the effect of short-duration sub-maximal cycling. Musculoskeletal Disorders. 2003;4(1):14.
57. Goldberg A, Casby A, Wasielewski M. Minimum detectable change for single-leg-stance-time in older adults. Gait and Posture. 2011;33(4):737.
58. Haupstein T. Visual judgements of steadiness in one-legged stance: reliability and validity. Physiotherapy Research International. 2000;5(3):141-56.
59. Hertel J. Time-to-boundary measures of postural control during single leg quiet standing. Journal of Applied Biomechanics. 2006;22(1):67.
60. Dennis RJ, Finch CF, Elliot BC, Farhart PJ. The reliability of musculoskeletal screening tests used in cricket. Physical Therapy in Sport. 2008;9:9.
61. Wallace RF. Rates of ankle and foot injuries in active-duty US Army soldiers, 2000-2006. Military Medicine. 2011;176(3):283.
62. Wilkinson DM. Injuries and injury risk factors among British army infantry soldiers during predeployment training. Injury Prevention. 2011;17(6):381-7.
63. Larsen K. Can custom-made biomechanic shoe orthoses prevent problems in the back and lower extremities? A randomized, controlled intervention trial of 146 military conscripts. Journal of Manipulative and Physiological Therapeutics. 2002;25(5):326.
64. Birrell SA, Hooper RH, Haslam RA. The effect of military load carriage on ground reaction forces. Gait and Posture. 2007;26(4):611-4.

65. Blacker SD, Wilkinson DM, Bilzon JL, Rayson MP. Risk factors for training injuries amoung British Army recruits. Military Medicine. 2008;173(3):278-86.
66. Hauret KG. Musculoskeletal Injuries:: Description of an under-recognized injury problem among military personnel. American Journal of Preventive Medicine. 2010;38(1):S61.
67. Gabbe BJ. Reliability of common lower extremity musculoskeletal screening tests. Physical Therapy in Sport. 2004;5(2):90-7.
68. Haight HJ. Measuring standing hindfoot alignment: Reliability of goniometric and visual measurements. Archives of Physical Medicine and Rehabilitation. 2005;86(3):571-5.
69. Bahr R, Krosshaug T. Understanding injury mechanisms: a key component of preventing injuries in sport. British Journal of Sports Medicine. 2005;39(6):324-9.
70. Birrell SA. The effect of load distribution within military load carriage systems on the kinetics of human gait. Applied Ergonomics. 2010;41(4):585.
71. Mahaudens P, Detrembleur C, Mousny M, Banse X. Gait in adolescent idiopathic scoliosis: energy cost analysis. European Spine Journal. 2009;18(8):1160-8.
72. Stolwijk NM. Plantar pressure changes after long-distance walking. Medicine and Science in Sports and Exercise. 2010;42(12):2264-72.
73. Keller M, Grossman J, Caron M, Mendicino R. Lateral ankle instability and the brostrom-gould procedure. The Journal of Foot and Ankle Surgery. 1996;35(6):513-20.
74. Callaghan MJ, Selfe J, Bagley PJ, Oldham JA. The effects of patellar taping on knee joint proprioception. Journal of Athletic Training. 2002;31(1):19-24.
75. Kerkhoffs G. A measurement device for anterior laxity of the ankle joint complex. Clinical Biomechanics (Bristol). 2005;20(2):218.
76. Dayakidis MK. Ground reaction force data in functional ankle instability during two cutting movements. Clinical Biomechanics (Bristol). 2006;21(4):405-11.
77. Zhang Y, Anderson S. Investigation of energy expenditure for exoskeleton walking: a case study. International Journal of Biomechatronics and Biomedical Robotics. 2012;2(1):26-31.
78. Aeillo D, Silver J, Gawne A. Gait, orthoses, footwear, and assistive devices. Philidelphia: Hanley and Belfus; 2004. 219-38 p.
79. O'Leary K, Vorpahl KA, Heiderscheit B. Effect of cushioned insoles on impact forces during running. Journal of the American Podiatric Medical Association. 2008;98(1):36-41.
80. Baxter ML, Baxter DG. Anthropometric characteristics of feet of soldiers in the New Zealand army. Military Medicine. 2011;176(4):438-45.
81. Laughton CA. Effect of strike pattern and orthotic intervention on tibial shock during running. Journal of Applied Biomechanics. 2003;19(2):153.
82. Razeghi M. Biomechanical analysis of the effect of orthotic shoe inserts: a review of the literature. Sports Medicine (Auckland). 2000;29(6):425.
83. AstrÃm M. Alignment and joint motion in the normal foot. The Journal of Orthopaedic and Sports Physical Therapy. 1995;22(5):216-22.
84. Briggs PJ. The structure and function of the foot in relation to injury. Current Orthopaedics. 2005;19(2):85-93.
85. Chang CH. Dynamic pedobarograph in evaluation of varus and valgus foot deformities. Journal of Pediatric Orthopaedics. 2002;22(6):813.
86. Collins N, Bisset LM, McPoil T, Vicenzino B. Foot orthoses in lower limb overuse conditions: A systematic review and meta-analysis. Foot and Ankle International. 2006;28(3):396-412.
87. Carr K. How can you help athletes prevent and treat shin splints? The Journal of Family Practice. 2008;57(6):406.

88. Hreljac A. Evaluation of lower extremity overuse injury potential in runners. Medicine and Science in Sports and Exercise. 2000;32(9):1635.
89. Johanson M, Baer J, Hovermale H, Phouthavong P. Subtalar joint position during gastrocnemius stretching and ankle dorsiflexion range of motion. Journal of Athletic Training. 2008;43(2):172.
90. Williams I, Dorsey S. Arch structure and injury patterns in runners. Clinical Biomechanics (Bristol). 2001;16(4):341-7.
91. Dixon SJ. Influence of a commercially available orthotic device on rearfoot eversion and vertical ground reaction force when running in military footwear. Military Medicine. 2007;172(4):446.
92. Fallat L, Grimm D, Saracco D. Sprained ankle syndrome: Prevalence and analysis of 639 acute injuries. The Journal of Foot and Ankle Surgery. 1998;37(4):280-5.
93. Lephart SM. The role of proprioception in the management and rehabilitation of athletic injuries. The American Journal of Sports Medicine. 1997;25(1):130-7.
94. Spanos S. The effect of taping on the proprioception of the ankle in a non-weight bearing position, amongst injured athletes. Foot (Edinburgh, Scotland). 2008;18(1):25-33.
95. Chen B. Ligament injuries of the ankle joint: a review of the Asian literature. Current Opinion in Orthopaedics. 1999;10(3):206.
96. Wikstrom E, Tillman M, Chmielewski T, Cauraugh J, Naugle K, Borsa P. Dynamic postural control but not mechanical stability differs among those with and without chronic ankle instability. Scandinavian Journal of Medicine and Science in Sports. 2010;20(1):137-44.
97. Grier TL. Footwear in the United States Army Band: Injury incidence and risk factors associated with foot pain. Foot (Edinburgh, Scotland). 2011.
98. Delisa J, Gans B, Walsh N, Robinson L. Physical Medicine and Rehabilitation. 5 ed. Frontera W, editor. Philadelphia: Lippincott, Williams and Wilkins; 2010.
99. Hughes T. The effects of proprioceptive exercise and taping on proprioception in subjects with functional ankle instability: a review of the literature. Physical Therapy in Sport. 2008;9(3):136.
100. Zazulak BT. Deficits in neuromuscular control of the trunk predict knee injury risk: A prospective biomechanical-epidemiologic study. The American Journal of Sports Medicine. 2007;35(7):1123-30.
101. Simoneau GG. Changes in ankle joint proprioception resulting from strips of athletic tape applied over the skin. Journal of Athletic Training. 1997;32(2):141.
102. De Jong A. Performance in different proprioceptive tests does not correlate in ankles with recurrent sprain. Archives of Physical Medicine and Rehabilitation. 2005;86(11):2101-5.
103. Deshpande N. Reliability and validity of ankle proprioceptive measures. Archives of Physical Medicine and Rehabilitation. 2003;84(6):883.
104. Eils E, Schroter R, Schroder M, Gerss J, Rosenbaum D. Multistation propriocetive exercise program prevents ankle injuries in basketball. Medicine and Science inSports and Exercise. 2010;42(11):2098-215.
105. Munn J. Evidence of sensorimotor deficits in functional ankle instability: a systematic review with meta-analysis. Journal of Science and Medicine in Sport. 2010;13(1):2.
106. Papadopoulos E. The role of ankle bracing in injury prevention, athletic performance and neuromuscular control: a review of the literature. Foot (Edinburgh, Scotland). 2005;15(1):1-6.
107. Teyhen DS. Awareness and compliance with recommended running shoe guidelines among US Army soldiers. Military Medicine. 2010;175(11):847.
108. Dixon SJ, Waterworth C, Smith C, House C. Military boots with new and degraded insoles. Medicine and Science in Sports and Exercise. 2003;35(3):472-9.

109. Rauh MJ, Macera CA, Trone DW, Shaffer RA, Brodine SK. The epidemiology of stress fractures and lower extremity overuse injury in female recruits. Medicine and Science in Sports and Exercise 2006;38(9):1571-7.
110. Constantini N, Mann G, Nyska M, Mei-dan O, Even A, Kahn G, editors. Risk factors for stress fractures and other orthopaedic injuries in female infantry recruits. Rendez-vouz Medicine Symposium; 2004 March; Vancouver.
111. Knapik JJ. Injury reduction effectiveness of assigning running shoes based on plantar shape in Marine corps basic training. The American Journal of Sports Medicine. 2010;38(9):1759-67.
112. Reid P, Robson B. Understanding health inequalities. In: Mulholland M, editor. The State of Maori Health. Auckland: Reed; 2006.
113. Ellison-Loschmann L, Pearce N. Improving access to health care among New Zealands Maori population. American Journal of Public Health. 2006;96:612-7.
114. Reid P, Cram F. Connecting health, people and country in Aotearoa New Zealand. In: Dew K, Davis IS, editors. Health and Society in Aotearoa New Zealand. Auckland: Otago University Press; 2005.
115. Raissi GR, Cherati ADS, Mansoori KD, Razi MD. The relationship between lower extremity alignment and medial tibial stress syndrome among non-professional athletes. Sports Medicine, Arthroscopy, Rehabilitation, Therapy and Technology. 2009;1(11):8.
116. Hegedus EJ, Cook C, Fiander C, Wright A. Meaures of arch height and their relationship to pain and dysfunction in people with lower limb impairments. Physiotherapy Research International. 2009;15:7.
117. Riskowski JL, Dufour AB, Hagedorn TJ, HIllstrom HJ, Casey VA, Hannan MT. Assoications of foot posture and function to lower extremity pain: Results from a population based foot study. Arthritis Care and Research. 2013;65(11):1804-12.
118. Leung AKL, Cheng JCY, Mak AFT. Orthotic design and foot impression procedures to control foot alignment. Prosthetics and Orthotics International. 2004;28(3):254.
119. Willems TM, Clercq D, Delbaere K, Vanderstraeten G, De Cock A, Witvrouw E. A prospective study of gait related risk factors for exercise-related lower leg pain. Gait and Posture. 2006;23(1):91-8.
120. Bennett JE, Reinking MF, Pluemer B, Pentel A, Seaton M, Killian C. Factors contributing to the development of medial tibial stress syndrome in high school runners. Journal of Orthopaedic and Sports Physical Therapy. 2001;31(9):504-10.
121. Leggat P. Military training and musculoskeletal disorders. Journal of Musculoskeletal Pain. 2007;15(2):25-32.
122. Gillespie WJ, Grant I. Interventions for preventing and treating stress fractures and stress reactios of bone of the lower limb. Cochrane Library. 2005;1:24.
123. Sahar T, Cohen M, Ne'eman V, Kandel L, Odebiyi D, Lev I, et al. Insoles for prevention and treatment of back pain. Cochrane Library. 2010;3:24.
124. Canham-Chervak M. A systematic process to prioritize prevention activities: Sustaining progress toward the reduction of military injuries. American Journal of Preventive Medicine. 2010;38(1):S11.
125. Jones BH. Prevention of lower extremity stress fractures in athletes and soldiers: A systematic review. Epidemiologic Reviews. 2002;24(2):228-47.
126. De Cock A, Clercq D, Willems TM, Witvrouw E. Temporal characteristics of foot roll-over during barefoot jogging: reference data for young adults. Gait and Posture. 2005;21(4):432-9.
127. Chen H. Influence of sensory input on plantar pressure distribution. Clinical Biomechanics (Bristol). 1995;10(5):271.

128. Jones BH. Physical training and exercise-related injuries: surveillance, research and injury prevention in military populations. Sports Medicine (Auckland). 1999;27(2):111.
129. Baxter M. Prevention of overuse injuries in the lower limb and lower back through orthotic intervention: an overview of current challenges and future directions. Physical Therapy Reviews. 2010;15(5):367.
130. Snyder RA. Does shoe insole modification prevent stress fractures? A systematic review. HSS Journal. 2009;5(2):92-8.
131. Landorf K, Keenan A-M. Do foot orthosis prevent injury? A systematic review. Journal of Science and Medicine in Sport. 2010;12:81.
132. Brushoj C. Prevention of overuse injuries by a concurrent exercise program in subjects exposed to an increase in training load: A randomized controlled trial of 1020 army recruits. The American Journal of Sports Medicine. 2008;36(4):663-70.
133. Herbert RD. Effects of stretching before and after exercising on muscle soreness and risk of injury: systematic review. British Medical Journal (Clinical research ed). 2002;325(7362):19.
134. Landorf K. Shock absorbing insoles and lower limb injury. Journal of the Royal Society of Medicine. 2006;99(1):3-4. Epub 1.
135. Gribble PA, Hertel J, Plisky P. Using the Star Excursion Balance Test to assess dynamic postural-control deficits and outcomes in lower extremity injury: A literature and systematic review. Journal of Athletic Training. 2012;47(3):339-57.
136. Munro AG, Herrington LC. Between-session reliability of four hop tests and the agility T-test. Journal of Strength and Conditioning Research. 2011;25(5):1470-7.
137. Meijne W, Neerbos K, Aufdemkampe G, Van Der Wurff P. Intraexaminer and interexaminer reliability of the Gillet test. Journal of Manipulative and Physiological Therapeutics. 1999;22(1):4-9.
138. Murphy DR. Interexaminer reliability of the hip extension test for suspected impaired motor control of the lumbar spine. Journal of Manipulative and Physiological Therapeutics. 2006;29(5):374.
139. Guggenmoos-Holzmann I. The meaning of kappa: probabilistic concepts of reliability and validity revisited. Journal of Clinical Epidemiology. 1996;49(7):775.
140. Vincent-Smith B. Inter-examiner and intra-examiner reliability of the standing flexion test. Manual Therapy. 1999;4(2):87.
141. Dean-Brown J. What is Validity? JALT Testing and Evaluation. 2000;4(2):8-12.
142. Robinson I, Dyson R, Halson-Brown S. Reliability of clinical and radiographic measurement of rearfoot alignment in a patient population. The Foot. 2001;11(1):2-9.
143. Saltzman CL, El-Khoury GY. The hindfoot alignment view. Foot and Ankle International. 1995;16(9):572-6.
144. Eechaute C, Bautmans I, De Hertogh W, Vaes P. The multiple hop test: a discriminative or evaluative instrument for chronic ankle instability? Clinical journal of sport medicine. 2012;22(3):228-33.
145. Finnoff JT, Peterson VJ, Hollman JH, Smith J. Intrarater and interrater reliability of the Balance Error Scoring System (BESS). Physical Medicine and Rehabilitation. 2009;1(1):50-4.
146. Trojian TH, McKeag DB. Single leg balance test to identify risk of ankle sprains British Journal of Sports Medicine. 2006;40(7):610-3.
147. Ayala F, De Baranda PS, De Ste Croix M, Santonja F. Reproducibility and concurrent validity of hip joint angle test for estimating hamstring flexibility in recreationally active young men. Journal of Strength and Conditioning Research. 2012;26(9):2372-82.
148. Demeritt KM, Shultz SJ, Docherty CL. CAI does not affect lower extremity functional performance. Journal of Athletic Training. 2002;37(5):507.

149. Eechaute C, Vaes P, Duquet W. The dynamic postural control is impaired in patients with chronic ankle instability: reliability and validity of the multiple hop test. Clinical Journal of Sport Medicine. 2009;19(2):107-14.
150. Frigg A, Nigg B, Davis E, Pederson B, Valderrabano V. Does alignment in the hindfoot radiograph influence dynamic foot-floor pressures in ankle and tibiotalocalcaneal fusion? Clinical Orthopaedics and Related Research. 2010;468(12):3362-70.
151. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet (British edition). 1986;February(8):4.
152. Whatman C. Physiotherapist agreement when visually rating movement quality during lower extremity functional screening tests. Physical Therapy in Sport. 2012;13(2):87-96.
153. Trojian TH. Single leg balance test to identify risk of ankle sprains British Journal of Sports Medicine. 2006;40(7):610-3.
154. Zazulak BT. The effects of core proprioception on knee injury: A prospective biomechanical-epidemiological study. The American Journal of Sports Medicine. 2006;35(3):368-73.
155. Hrysomallis C. Relationship between balance ability, training and sports injury risk. Sports Medicine (Auckland). 2007;37(6):547.
156. Laslett M, Williams M. The reliability of selected pain provaction tests for sacroilliac joint pathology. Spine. 1994;19:1243-9.
157. Webb P, Pain C. Essential Epidemiology. 2 ed. New York: Cambridge University Press; 2011.
158. Buck P. The optimum position of arthrodesis of the ankle. A gait study of the knee and ankle. Journal of Bone and Joint Surgery (American Volume). 1987;69(7):1052-62.
159. McPoil TG. Relationship between three static angles of the rearfoot and the pattern of rearfoot motion during walking. The Journal of Orthopaedic and Sports Physical Therapy. 1996;23(6):370-5.
160. Schneiders AG. Functional movement screen; normative values for a young, active population. International Journal of Sports Physical Therapy. 2011;6(2):75.
161. Rome K, Handoll H, Ashford R. Interventions for preventing and treating stress fractures and stress reactions of bone of the lower limbs in young adults. Cochrane Library. 2009(1):51.
162. Murphy DF, Connolly DAJ, Beynnon BD. Risk factors for lower extremity injury: a review of the literature. British Journal of Sports Medicine. 2003;37(1):13-29.
163. Catling-Paull C, Johnston R, Ryan C, Foureur MJ, Homer CS. Clinical interventions that increase the uptake and success of vaginal birth after caesarean section: a systematic review. Journal of Advanced Nursing. 2011;67(8):1646-61.
164. Orava S. Stress fractures. British journal of sports medicine. 1980;14(1):40.
165. Khan KM, Fuller PJ, Brukner PD, Kearney C, Burry HC. Outcome of conservative and surgical management of navicular stress fracture in athletes Eighty-six cases proven with computerized tomography. The American Journal of Sports Medicine. 1992;20(6):657-66.
166. Sherrington C, Herbert R, Maher C, Moseley A. PEDro. A database of randomized trials and systematic reviews in physiotherapy. Manual Therapy. 2000;5(4):223-6.
167. Finestone A, Giladi M, Elad H, Salmon A, Mendel D, Milgrom C. A randomized clinical trial of the effect of custom biomechanical shoe orthotics on the incidence of stress fractures. Clinical Orthopedics and Related Research. 1999;360:182-90.
168. Gardner Jr LI, Dziados JE, Jones BH, Brundage JF, Harris J, Sullivan R, et al. Prevention of lower extremity stress fractures: a controlled trial of a shock absorbent insole. American Journal of Public health. 1988;78(12):1563-7.

169. Sherman R, Karstetter K. Prevention of lower limb pain in soldiers using shock absorbing orthotic inserts. Journal of the American Podiatric Medical Association. 1996;83(3):6.
170. Milgrom C, Giladi M, Stein M, Kashtan H, Margulies J, Chisin R, et al. Stress fractures in military recruits. A prospective study showing an unusually high incidence. Journal of Bone and Joint Surgery, British Volume. 1985;67(5):732-5.
171. Milgrom C, Finestone A, Lubovsky O, Zin D, Lahad A. A controlled randomized study of the effect of training with orthoses on the incidence of weight bearing induced back pain among infantry recruits. Spine. 2005;30(3):272-5.
172. Mattila VM. Orthotic insoles do not prevent physical stress-induced low back pain. European spine journal. 2011;20(1):100-4.
173. Simkin A, Leichter I, Giladi M, Stein M, Milgrom C. Combined effect of foot arch structure and an orthotic device on stress fractures. Foot and Ankle International. 1989;10(1):25-9.
174. Warren BL. Determining predictor variables for running-related pain. Physical Therapy. 1988;68(5):647.
175. Serfontein JH. A prediction model for the prevention of soccer injuries amongst youth players: North-West University, Potchefstroom Campus; 2009.
176. Yamazaki J. Differences in kinematics of single leg squatting between anterior cruciate ligament-injured patients and healthy controls. Knee Surgery, Sports Traumatology, Arthroscopy. 2010;18(1):56-63.
177. Niemuth PE. Hip muscle weakness and overuse injuries in recreational runners. Clinical Journal of Sports Medicine. 2005;15(1):14.
178. Yamamoto H. Subtalar instability following lateral ligament injuries of the ankle. Injury. 1998;29(4):265.
179. Donoghue OA. Functional data analysis of running kinematics in chronic achilles tendon injury. Medicine and Science in Sports and Exercise. 2008;40(7):1323-35.
180. Willems TM. Gait-related risk factors for exercise-related lower-leg pain during shod running. Medicine and Science in Sports and Exercise. 2007;39(2):330-9.
181. Cuğ M. The effect of instability training on knee joint proprioception and core strength. Journal of Sports Science and Medicine. 2012;11:468.
182. Bonnel F. Chronic ankle instability: biomechanics and pathomechanics of ligaments injury and associated lesions. Orthopaedics and Traumatology, Surgery and Research. 2010;96(4):424.
183. Bullinger KL. Permanent central synaptic disconnection of proprioceptors after nerve injury and regeneration. II. Loss of functional connectivity with motoneurons. Journal of Neurophysiology. 2011;106(5):2471-85.
184. Schneiders AG. Normative values for three clinical measures of motor performance used in the neurological assessment of sports concussion. Journal of Science and Medicine in Sport. 2010;13(2):196-201.
185. Aruin AS. Anticipatory postural adjustments in conditions of postural instability. Electroencephalography and Clinical Neurophysiology. 1998;109(4):350.
186. Lee AJY. Twelve-week biomechanical ankle platform system training on postural stability and ankle proprioception in subjects with unilateral functional ankle instability. Clinical Biomechanics (Bristol). 2008;23(8):1065.
187. Jonsson E. One-leg stance in healthy young and elderly adults: a measure of postural steadiness? Clinical Biomechanics (Bristol). 2004;19(7):688-94.
188. Hoffman M. Unilateral postural control of the functionally dominant and nondominant extremities of healthy subjects. Journal of Athletic Training. 1998;33(4):319.

189. Warren J. Repeated single-limb postural stability testing elicits a practice effect. Physical Therapy in Sport. 2006;7(4):185.
190. Springer BA. Normative values for the unipedal stance test with eyes open and closed. Journal of Geriatric Physical Therapy. 2007;30(1):8.
191. Tyson S, DeSouza L. Brunel standing balance tests. Physiotherapy. 2002;88(7):430.
192. Ross SE. Examination of static and dynamic postural stability in individuals with functionally stable and unstable ankles. Clinical Journal of Sport Medicine. 2004;14(6):332.
193. Choy NL. Changes in postural stability in women aged 20 to 80 years. The Journals of Gerontology. 2003;58(6):M525-M30.
194. Fuller E. The Windlass mechanism of the foot. A mechanical model to explain pathology. Journal of the American Podiatric Medical Association. 2000;90:35-46.
195. Kirby KA. Methods for determination of positional variations in the subtalar joint axis. Journal of the American Podiatric Medical Association. 1987;77(5):228-34.
196. Zinder SM. Validity and reliability of a new in vivo ankle stiffness measurement device. Journal of Biomechanics. 2007;40(2):463.
197. Lamm BM, Paley D, Herzenberg JE. Gastrocnemius Soleus Recession. Journal of the American Podiatric Medical Association. 2005;95(1):18-25.
198. Nadeau S. Plantarflexor weakness as a limiting factor of gait speed in stroke subjects and the compensating role of hip flexors. Clinical Biomechanics (Bristol). 1999;14(2):125-35.
199. Olree KS. Fundamental patterns of bilateral muscle activity in human locomotion. Biological Cybernetics. 1995;73(5):409-14.
200. Hébert-Losier K. Analysis of knee flexion angles during 2 clinical versions of the heel raise test to assess soleus and gastrocnemius function. The Journal of Orthopaedic and Sports Physical Therapy. 2011;41(7):505.
201. Gottschall JS. Energy cost and muscular activity required for propulsion during walking. Journal of Applied Physiology. 2003;94(5):1766.
202. Valovich TC. Repeat administration elicits a practice effect with the Balance Error Scoring System but not with the Standardized Assessment of Concussion in high school athletes. Journal of Athletic Training. 2003;38(1):51.
203. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Medical Research Methodology. 2010;10(1):67.
204. Justis RT, Kreigsmann B. The feasibility study as a tool for venture analysis. Business Journal of Small Business Management. 1979;17(1):35-42.
205. Georgakellos DA, Macris AM. Application of the semantic learning approach in the feasibility studies preparation training process. Information Systems Management. 2009;26(3):231-40.
206. Young G. Feasibility studies. Appraisal Journal. 1970;38(3):376-83.
207. Viehweger E, Pfund LZ, Hélix M, Rohon M-A, Jacquemier M, Scavarda D, et al. Influence of clinical and gait analysis experience on reliability of observational gait analysis (Edinburgh Gait Score Reliability). Annals of Physical and Rehabilitation Medicine. 2010;53(9):535-46.
208. Tian ZZ, Kyte MD, Messer CJ. Parallax error in video-image systems. Journal of Transportation Engineering. 2002;128(3):218-23.
209. Garzia RP. Foundations of Binocular Vision: A Clinical Perspective. New York: McGraw Hill; 2000.
210. Simpson JA, Weinar ESC. The Oxford English Dictionary. London: Clarendon Press; 2002.

211. Reimann BL, Myers JB, Lephart SM. Comparison of the ankle, knee, hip and truck corrective action shown during single-leg stance on firm, foam and multiaxial surfaces. Archives of Physical Medicine and Rehabilitation. 2003;80:6.
212. Brosseau-Laird PE, Savalei V. When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. Psychological Methods. 2012;17(3):20.
213. Hastings JD, Fanucchi ER, Burns SP. Wheelchair configuration and postural alignment in persons with spinal cord injury. Archives of physical medicine and rehabilitation. 2003;84(4):528-34.
214. Caron O, Faure B, Brenière Y. Estimating the centre of gravity of the body on the basis of the centre of pressure in standing posture. Journal of Biomechanics. 1997;30(11):1169-71.
215. Stata Statistical Software Release 7.0: Programming: Stata Corporation; 2001.
216. Sim J, Wright CC. The kappa statistic in reliabiltiy studies: use, interpretation and sample size requirements. Physical Therapy. 2005;85:257-68.
217. Stochkendahl MJ, Christensen HW, Hartvigsen J, Vach W, Haas M, Hestbaek L, et al. Manual examination of the spine: a systematic critical literature review of reproducibility. Journal of Manipulative and Physiological Therapeutics. 2006;29(6):475.
218. Sports Medicine New Zealand. 2013 [cited 2013 1st November]; Available from: http://www.sportsmedicine.co.nz/node/9.
219. Donner A. Sample size requirements for reliability studies. Statistics in Medicine. 1987;6(4):441-8.
220. Randolph JJ. Online Kappa Calculator. 2008 [cited 2013 April 8th]; Available from: http://justus.randolph.name/kappa.
221. Thomson CE. An investigation into the reliability of the valgus index and its validity as a clinical measurement. Foot. 1994;4:7.
222. Pownall PJ, Moran RW, Stewart AM. Consistency of standing and seated posture of asymptomatic male adults over a one-week interval: a digital camera analysis of multiple landmarks. International Journal of Osteopathy and Medicine. 2008;11:8.
223. Watson AWS, MacDonncha C. A reliable technique for the assessment of posture; assessment criteria for aspects of posture. Journal of Sports Medicine and Physical Fitness. 2000;40:10.
224. Bovens AM, van Baak MA, Vrencken JG, Wijnen JA, Verstappen FT. Variability and reliability of joint measurements. The American Journal of Sports Medicine. 1990;18(1):58-63.
225. Hertel J, Olmsted-Kramer LC. Deficits in time-to-boundary measures of postural control with chronic ankle instability. Gait and Posture. 2007;25(1):33-9.
226. Khan FM, Khan F. The physics of radiation therapy: Lippincott Williams & Wilkins Philadelphia; 2003.
227. Nosewicz TL, Knupp M, Bolliger L, Hintermann B. The reliability and validity of radiographic measurements for determining the three-dimensional position of the talus in varus and valgus osteoarthritic ankles. Journal of the International Skeletal Society. 2012;10:11.
228. Milosavljevic S, McBride DI, Bagheri N, Vasiljev RM, Mani R, Carman AB, et al. Exposure to whole body vibration and mechanical shock: A field study of quad bike use in agriculture. Annals of Occupational Hygiene. 2011;55(3):10.
229. Silbernagel KG, Nilsson-Helander K, Thomee R, Ericksson B, Karlsson J. A new measurement of heel-raise endurance with the ability to detect functional deficits in patients with achilles tendon rupture. Knee Surgery, Sports Traumatology, Arthroscopy [Internet]. 2009:[7 p.].

230. Veilleux LN, Rauch F. Reproducibility of jumping mechanography in healthy children and adults. Journal of Musculoskeletal and Neuronal Interactions. 2010;10(4):11.
231. Zammit GV, Menz HB, Munteanu SE. Research reliability of the TekScan MatScan® system for the measurement of plantar forces and pressures during barefoot level walking in healthy adults. Journal of Foot and Ankle Research. 2010;3(11):9.
232. Wiles C, Newcombe R, Fuller K, Jones A, Price M. Use of videotape to assess mobility in a controlled randomized crossover trial of physiotherapy in chronic multiple sclerosis. Clinical Rehabilitation. 2003;17(3):256-63.

18 Chapter 6, Appendix 1

VOLUNTEERS NEEDED!

Establishing the reliability and validity of the FootFAST lower limb screening protocol for overuse injury: A feasibility study.

Clinically trained and experienced physiotherapists are needed as assessors in the study titled above. If you are current practising as a physiotherapist, and have had a minimum of three years of clinical experience, you are eligible for inclusion.

The aim of the study is to investigate feasibility of a study to determine the reliability and validity of the three visual assessments of function which comprise the FootFAST tool. Research such as this allows clinicians to make informed choices and justify their practises.

The time commitment involved in the study is between one to two hours, on three separate occasions. In short, you will be required to assess ten 'patients' using the FootFAST tool. This same assessment will be repeated twice (for a total of three times). There is no requirement for you to travel as video footage of the 'patients' will be provided to you on a DVD, along with hard copies of the scoring sheets.

If you are interested in volunteering to be part of the study, please contact the researcher using any of the details provided below. Information sheets for assessors and the consent forms can be provided on request.

This study has been approved by the University of Otago Human Ethics Committee.

Marian Baxter
marian.baxter@otago.ac.nz
027 378 2757
03 4544292
Office: Room 121, First Floor, Adams Building, Frederick Street.

19 Chapter 6, Appendix 2

INFORMATION SHEET FOR PARTICIPANTS

Establishing the reliability and validity of the FootFAST lower limb screening protocol for overuse injury: A feasibility study.

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether or not to participate. If you decide to participate we thank you. If you decide not to take part there will be no disadvantage to you and we thank you for considering our request.

What is the Aim of the Project?

The aim is to identify the feasibility of testing the reliability and validity of the screening protocol: FootFAST. FootFAST is a screening protocol which assesses risk of overuse injury in the lower limb.

The results of this study will be used to inform the future uses and development of the protocol.

What Type of Participants are being sought?

- If you are currently not suffering an injury which affects your ability to walk, and you are aged between 18 to 35 years of age, you are eligible to participate.

What will Participants be Asked to Do?

Should you agree to take part in this project, you will be asked to participate in four short activities. Please note that there will be digital recordings made during all of these activities in the form of pictures, video and pressure data. These activities are outlined as follows:
You will undertake three lower limb screening tasks. You will complete one set of the tasks in bare feet with shorts and comfortable clothing and with no markers placed on your lower limbs. The second set of tasks will be the same but you will stand on a force platform and have reflective markers placed on your skin for both foot loading and 3D movement analysis.

The tasks will involve:

a) Measurement of your hind foot (heel) posture in normal, two legged stance.
b) Balancing on one leg for as long as you can, with your eyes closed.
c) Raising your heel (heel raise) while standing on one leg with eyes open.
d) Walking across a foot scanning mat in bare feet.

The time commitment involved is approximately two hours and will take place in the University of Otago School of Physiotherapy motion analysis suite.

There is no added risk of discomfort to those who choose to participate. The inconvenience caused to you is minimal and primarily is related to the time commitment involved.

Please be aware that at any stage you may decide not to take part in the project without any disadvantage to yourself of any kind.

What Data or Information will be Collected and What Use will be Made of it?

- Personal identifying information collected for this experiment includes your name and date of birth, gender and ethnicity. This information will not be used in the analysis and will be destroyed at the completion of the project.
- The information collected will be viewed by the researchers and research student, namely, Associate Professor Stephan Milosavljevic, Associate Professor David McBride and Ms Marian Baxter. Additional members of Ms Baxter's supervisory team may also view the data to assist with the analysis.
- Pictures of your feet and video footage of you walking obtained during the four activities will also be collected. This data pertains to the results of the study, and therefore will be securely stored in such a way that only those mentioned above will be able to gain access to it. Data obtained as a result of the research will be retained for **at least 5 years** in secure storage. Any personal information about you [*such as name and date of birth*] may be destroyed at the completion of the research even though the data derived from the research will, in most cases, be kept for much longer or possibly indefinitely.
- The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve your anonymity.
- You have the right to view the data collected from yourself at any stage, you can contact any of the researchers using the information provided in order to do so.
- The results of the study will be made available to you upon request once data collection and analysis is completed. This is expected to be in September 2013.

Can Participants Change their Mind and Withdraw from the Project?

You may withdraw from participation in the project at any time and without any disadvantage to yourself of any kind.

What if Participants have any Questions?

If you have any questions about our project, either now or in the future, please feel free to contact one of the following persons:-

Marian L Baxter	Assoc Prof Stephan Milosavljevic	Assoc Prof David McBride
(03) 4762358	(03) 4797193	027 227 6546
marian.baxter@otago.ac.nz	stephan.milosavljevic@otago.ac.nz	david.mcbride@otago.ac.nz
Department of Preventive and Social Medicine University of Otago P.O. Box 913 Dunedin 9054 New Zealand	School of Physiotherapy University of Otago PO Box 56 Dunedin 9054 New Zealand	Department of Preventive and Social Medicine University of Otago P.O. Box 913 Dunedin 9054 New Zealand

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.

CONSENT FORM FOR PARTICIPANTS

Establishing the reliability and validity of the FootFAST lower limb screening protocol for overuse injury: A feasibility study.

I have read the Information Sheet concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I know that:-

1. My participation in the project is entirely voluntary;
2. I am free to withdraw from the project at any time without any disadvantage;
3. Personal identifying information in the form of my name and date of birth will be destroyed at the conclusion of the project but any raw data on which the results of the project depend will be retained in secure storage for at least five years;
4. I consent to having both pictures and video recordings made of my feet and movements and I understand that these will be used for the purposes of analysis of this research only, and that they will not be released and disseminated in any way.
5. The project has been designed to minimise the risk and inconvenience to myself, the participant.
6. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve my anonymity.
7. I understand what will be required should I choose to participate.

All questions that I have are answered to my satisfaction, and I am aware that I can ask any more questions should I have them in the future.

I agree to take part in this project.

.....

.....

(Signature of participant) (Date)

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.

20 Chapter 6, Appendix 3

INFORMATION SHEET FOR OBSERVERS

Establishing the reliability and validity of the FootFAST lower limb screening protocol for overuse injury: A feasibility study.

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether or not to participate. If you decide to participate we thank you. If you decide not to take part there will be no disadvantage to you and we thank you for considering our request.

What is the Aim of the Project?

The aim is to identify the feasibility of testing the reliability and validity of the screening protocol: FootFAST. FootFAST is a screening protocol which assesses risk of overuse injury in the lower limb.

The results of this study will be used to inform the future uses and development of the protocol.

What Type of Observers are being sought?

- If you are currently a postgraduate Physiotherapy student at the University of Otago, or Physiotherapy clinical teaching staff, you are eligible.

What will Observers be Asked to Do?

Should you agree to take part in this project, you will be asked to participate in four short activities. Please note that the assessments you will be making will be outlined and detailed in full. All of the pictures and videos will be given to you in one DVD.

Part 1; You will be provided with ten pictures of hind feet. Assess the hind foot angle using the scoring system provided.

Part 2; You will be provided with a video of ten single limb stance tests, Assess these performances using the scoring sheet provided.

Part 3; You will be provided with ten videos of a forefoot raise. Assess these performances using the scoring sheet provided.

Following the instructional session which will outline the use of the protocol (approximately 20 minutes), the time commitment involved is approximately 15 minutes per participant, approximately 150 minutes in total. You will be asked to perform this same procedure on two separate occasions. This can take place wherever is convenient to you, as all that is required is a DVD player.

There is no added risk of discomfort to those who choose to participate. The inconvenience caused to you is minimal and primarily is related to the time commitment involved.
Please be aware that at any stage you may decide not to take part in the project without any disadvantage to yourself of any kind.

What Data or Information will be Collected and What Use will be Made of it?

- Personal identifying information collected for this experiment includes your name and address. This information will not be used in the analysis and will be destroyed at the completion of the project.
- The information collected will be viewed by the researchers and research student, namely, Associate Professor David McBride, Associate Professor Stephan Milosavljevic and Ms Marian Baxter. Additional members of Ms Baxter's supervisory team may also view the data to assist with the analysis.
- Data obtained from the images and video footage collected during the three activities will also be collected. The data pertains to the results of the study, and therefore will be securely stored in such a way that only those mentioned above will be able to gain access to it. Data obtained as a result of the research will be retained for **at least 5 years** in secure storage. Any personal information held on the participants [*such as name and date of birth*] may be destroyed at the completion of the research even though the data derived from the research will, in most cases, be kept for much longer or possibly indefinitely.
- The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve your anonymity.
- You have the right to view the data collected from yourself at any stage, you can contact any of the researchers using the information provided in order to do so.
- The results of the study will be made available to you upon request once data collection and analysis is completed. This is expected to be in September 2013.

Can Observers Change their Mind and Withdraw from the Project?

You may withdraw from participation in the project at any time and without any disadvantage to yourself of any kind.

What if I have any Questions?

If you have any questions about our project, either now or in the future, please feel free to contact one of the following persons:-

Marian L Baxter	Assoc Prof Stephan Milosavljevic	Assoc Professor David McBride
(03) 4762358	(03) 4797193	027 227 6546
marian.baxter@otago.ac.nz	stephan.milosavljevic@otago.ac.nz	david.mcbride@otago.ac.nz
Department of Preventive and Social Medicine University of Otago P.O. Box 913 Dunedin 9054 New Zealand	School of Physiotherapy University of Otago PO Box 56 Dunedin 9054 New Zealand	Department of Preventive and Social Medicine University of Otago P.O. Box 913 Dunedin 9054 New Zealand

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.

CONSENT FORM FOR OBSERVERS

Establishing the reliability and validity of the FootFAST lower limb screening protocol for overuse injury: A feasibility study.

I have read the Information Sheet concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I know that:-

1. My participation in the project is entirely voluntary;
2. I am free to withdraw from the project at any time without any disadvantage;
3. Personal identifying information in the form of my name and address will be destroyed at the conclusion of the project but any raw data on which the results of the project depend will be retained in secure storage for at least five years;
4. The project has been designed to minimise the risk and inconvenience to myself, the participant.
5. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve my anonymity.
6. I understand what will be required should I choose to participate.

All questions that I have are answered to my satisfaction, and I am aware that I can ask any more questions should I have them in the future.

I agree to take part in this project.

.....

.....

(Signature of participant) (Date)

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.

21 Chapter 6, Appendix 4

Thank you in advance for your time involved in completing this assessment.

I hope that all the instructions provided are comprehensive and simplistic, but if you are stuck or confused, please feel free to contact me:

marian.baxter@otago.ac.nz

Phone or text: 027 378 2757

Basic info:

You are provided with ten videos containing a heel raise test, ten containing a single leg stance and ten pictures of hindfeet, all in their relevant folders.

You are also provided with a scoring sheet and some instructions in each folder.

The default media player of windows will likely run the videos, but if you wish to use slow motion, or cycle through frame by frame, use a media player such as VLC.

MOST IMPORTANTLY:

Please judge the performance as outlined in the information provided. There are many criteria for judging the performance of these tests, but please only stick to the criteria outlined. Always assign the worst level of scoring you observe.*

*Some of the scoring continuums are on a likert scale, and they are continuous. For example, If a participant is observed to have a very stable during the majority of their single leg stance performance, but then has a perturbation of eversion, you may choose to place your score slightly towards ‘eversion’, but closer to ‘stable’.

When you are finished filling in all the scoring sheets from all three folders, please return them via email. The scoring sheets are the only documents I need returned.

NB: You can work through the assessments in any order you like and take as much time as you feel necessary.

Recording Sheet:
Assessment of Heel Raise

Observe the video of a heel raise performance, numbered one to ten. Indicate on the scale shown how good you think the performance is. The definition of each classification is given:

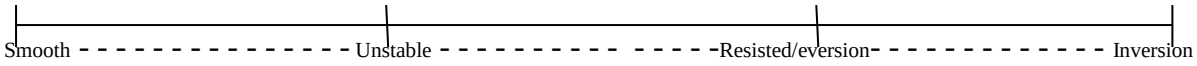
Smooth: The participant can move in a controlled manner into a high position on their toes.

Unstable: The foot may or may not be able to reach a good height, but the motion is not smooth.

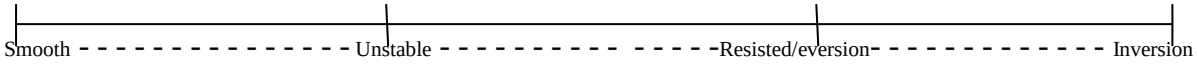
Resisted/ eversion: The foot struggles moving upwards, may even tend to fluctuate up and down, and there may be pronounced "falling inwards" where the foot twists in a toeing-in manner.

Inversion: The foot repeatedly "falls outwards" where the foot loses contact with the inside (medial) edge of ground surface, **and** height cannot be reached.

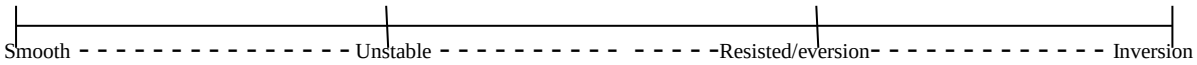
Video 1:



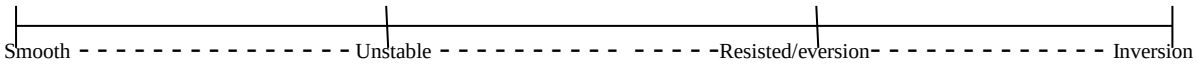
Video 2:



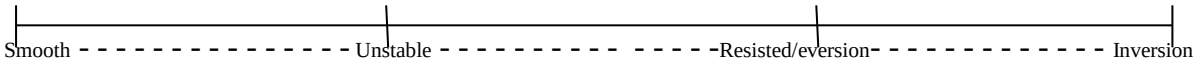
Video 3:



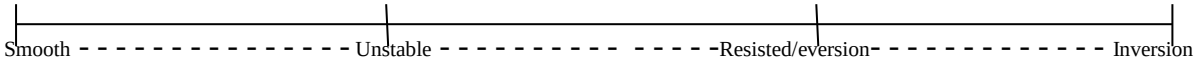
Video 4:



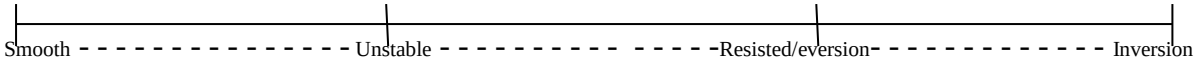
Video 5:



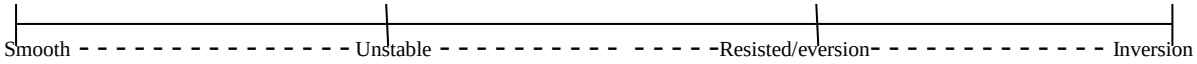
Video 6:



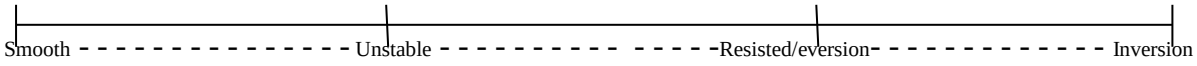
Video 7:



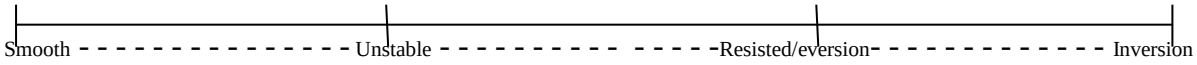
Video 8:



Video 9:



Video 10:

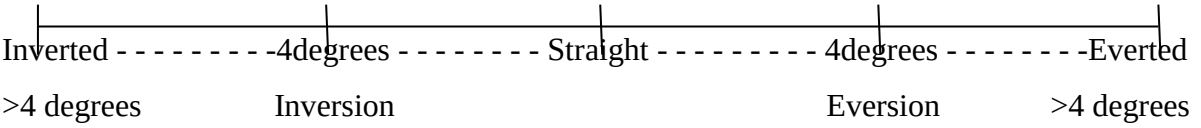


Recording Sheet:

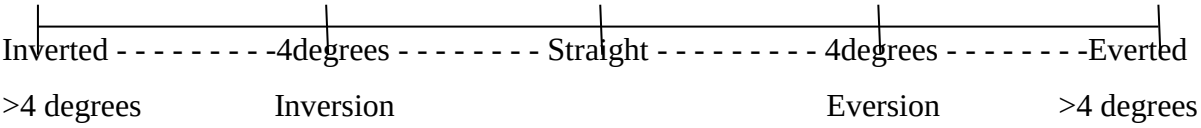
Assessment of Hind foot position

Observe the hindfoot shown in the pictures, numbered one to ten. Indicate on the scale shown where you think the position of the hindfoot is on the participants' **RIGHT FOOT**.

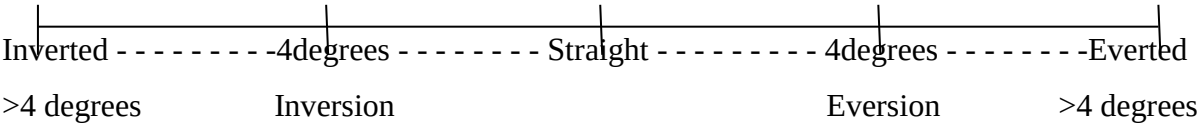
Picture 1:



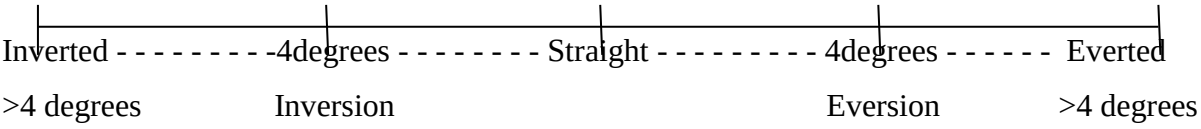
Picture 2:



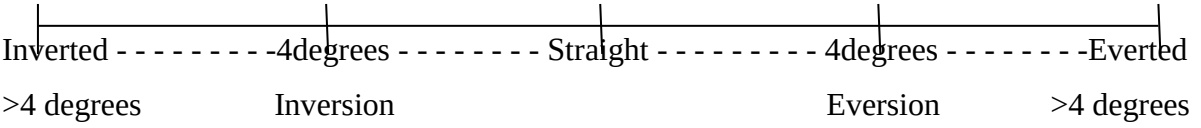
Picture 3:



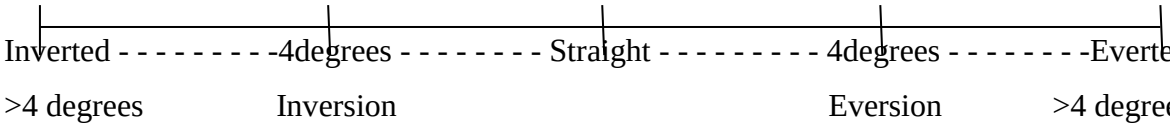
Picture 4:



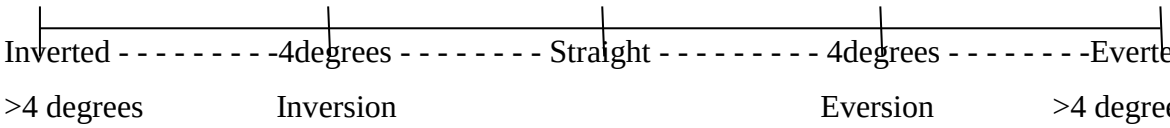
Picture 5:



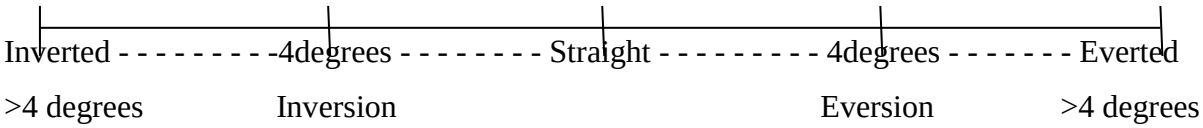
Picture 6:



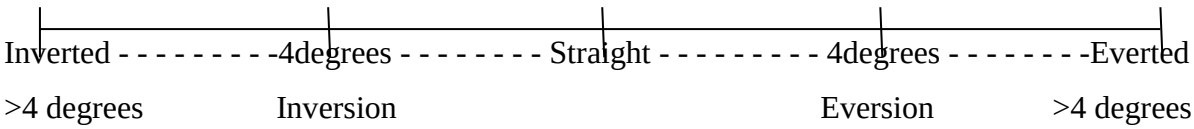
Picture 7:



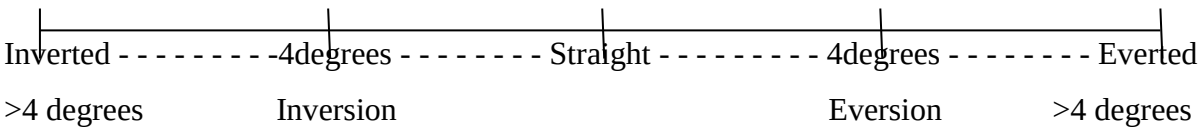
Picture 8:



Picture 9:



Picture 10:



Recording Sheet:
Assessment of Single leg stance

Observe the video of a single leg stance performance, numbered one to ten. Indicate on the scale shown how good you think the performance is. The definition of each classification is given:

Stable: The foot remains balanced, without perturbation or oscillations

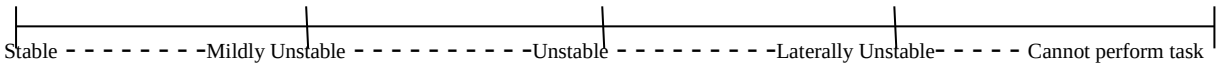
Mildly unstable: On occasion, the foot tends to oscillate or correct itself

Unstable: The foot is continuously oscillating and there may be a reduction in the ground-surface contact at times

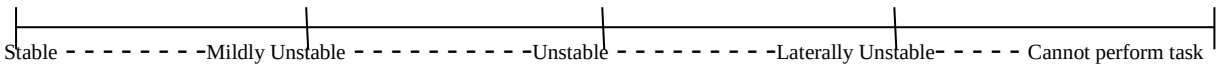
Laterally unstable: The foot continuously falls towards the lateral edge of the foot, losing ground contact on the medial aspect.

Cannot perform task: As indicated, if the participant is not balanced for at least 5 seconds.

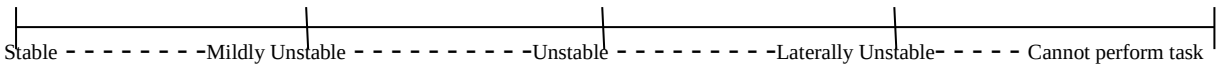
Video 1:



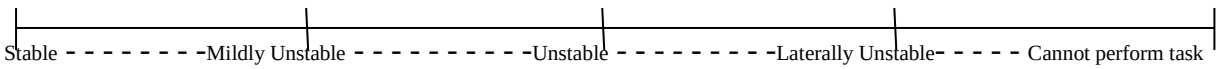
Video 2:



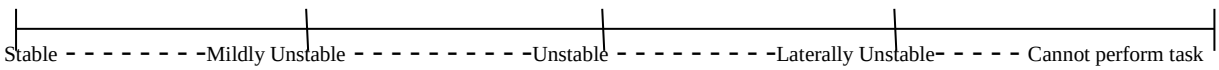
Video 3:



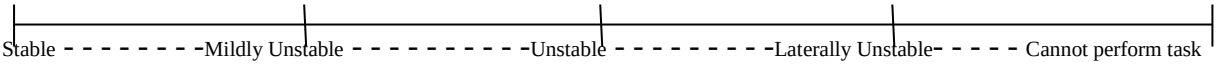
Video 4:



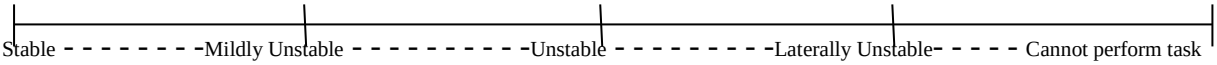
Video 5:



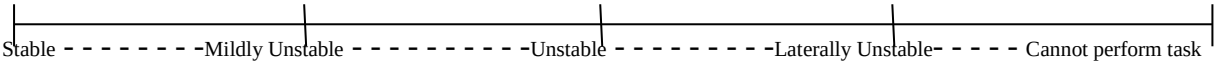
Video 6:



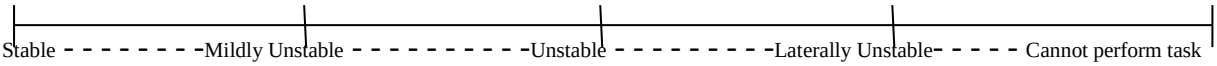
Video 7:



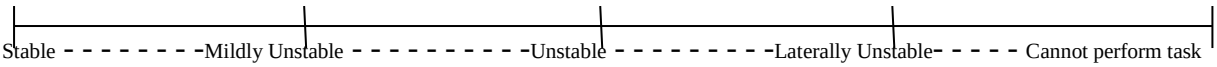
Video 8:



Video 9:



Video 10:



22 Chapter 7, Appendix 5

VOLUNTEERS NEEDED!

Establishing the reliability and validity of the FootFAST lower limb screening protocol for overuse injury.

Clinically trained and experienced *musculoskeletal physiotherapists*/*sports medicine practitioners/NZ army medics* are needed as assessors in the study titled above. If you have had a minimum of three years clinical experience, currently practising, and registered, you are eligible for inclusion.

The aim of the study is to investigate the reliability and validity of the three visual assessments of function which comprise the FootFAST tool. Research such as this allows clinicians to make informed choices and justify their practises.

The time commitment involved in the study is between two to three hours, on three separate occasions. In short, you will be required to assess eighteen 'patients' using the FootFAST tool. This same assessment will be repeated twice (for a total of three times). There is no requirement for you to travel as video footage of the 'patients' will be provided to you on a DVD, along with hard copies of the scoring sheets. All assessments must be completed by January 15th 2013.

If you are interested in volunteering to be part of the study, and can complete the assessments within the required time frame, please contact the researcher using any of the details provided below. Information sheets for assessors and the consent forms can be provided on request.

This study has been approved by the University of Otago Human Ethics Committee.

Marian Baxter

marian.baxter@otago.ac.nz

027 378 2757

03 4544292

Office: Room 121, First Floor, Adams Building, Frederick Street.

23 Chapter 7, Appendix 6

INFORMATION SHEET FOR PARTICIPANTS

Establishing the reliability and validity of the FootFAST lower limb screening protocol for overuse injury.

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether or not to participate. If you decide to participate we thank you. If you decide not to take part there will be no disadvantage to you and we thank you for considering our request.

What is the Aim of the Project?

The aim is to identify the reliability and validity of the screening protocol: FootFAST. FootFAST is a screening protocol which assesses risk of overuse injury in the lower limb. The results of this study will be used to inform the future uses and development of the protocol.

What Type of Participants are being sought?

If you are currently not suffering an injury which affects your ability to walk, and you are aged between 18 to 35 years of age, you are eligible to participate.

What will Participants be Asked to Do?

Should you agree to take part in this project, you will be asked to participate in four short activities. Please note that there will be digital recordings made during all of these activities in the form of pictures, video and pressure data. These activities are outlined as follows:

You will undertake three lower limb screening tasks. You will complete one set of the tasks in bare feet with shorts and comfortable clothing and with no markers placed on your lower limbs. The second set of tasks will be the same but you will stand on a force platform and have reflective markers placed on your skin for both foot loading and 3D movement analysis.

The tasks will involve:

- Measurement of your hind foot (heel) posture in normal two legged stance.
- Balancing on one leg for as long as you can, with your eyes closed.
- Raising your heel (heel raise) while standing on one leg with eyes open.
- Walking across a foot scanning mat in bare feet.

The time commitment involved is approximately two hours and will take place in the University of Otago School of Physiotherapy motion analysis suite.

There is no added risk of discomfort to those who choose to participate. The inconvenience caused to you is minimal and primarily is related to the time commitment involved.

Please be aware that at any stage you may decide not to take part in the project without any disadvantage to yourself of any kind.

What Data or Information will be Collected and What Use will be Made of it?

- Personal identifying information collected for this experiment includes your name and date of birth, gender and ethnicity. This information will not be used in the analysis and will be destroyed at the completion of the project.
- The information collected will be viewed by the researchers and research student, namely, Associate Professor Stephan Milosavljevic, Associate Professor David McBride and Ms Marian Baxter. Additional members of Ms Baxter's supervisory team may also view the data to assist with the analysis.
- Pictures of your feet and video footage of you walking obtained during the four activities will also be collected. This data pertains to the results of the study, and therefore will be securely stored in such a way that only those mentioned above will be able to gain access to it. Data obtained as a result of the research will be retained for **at least 5 years** in secure storage. Any personal information about you [*such as name and date of birth*] may be destroyed at the completion of the research even though the data derived from the research will, in most cases, be kept for much longer or possibly indefinitely.
- The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve your anonymity.
- You have the right to view the data collected from yourself at any stage, you can contact any of the researchers using the information provided in order to do so.
- The results of the study will be made available to you upon request once data collection and analysis is completed. This is expected to be in September 2013.

Can Participants Change their Mind and Withdraw from the Project?

You may withdraw from participation in the project at any time and without any disadvantage to yourself of any kind.

What if Participants have any Questions?

If you have any questions about our project, either now or in the future, please feel free to contact one of the following persons:-

Marian L Baxter	Assoc Prof Stephan Milosavljevic	Assoc Prof David McBride
(03) 4762358	(03) 4797193	027 227 6546
marian.baxter@otago.ac.nz	stephan.milosavljevic@otago.ac.nz	david.mcbride@otago.ac.nz
Department of Preventive and Social Medicine University of Otago P.O. Box 913 Dunedin 9054 New Zealand	School of Physiotherapy University of Otago PO Box 56 Dunedin 9054 New Zealand	Department of Preventive and Social Medicine University of Otago P.O. Box 913 Dunedin 9054 New Zealand

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.

CONSENT FORM FOR

PARTICIPANTS

Establishing the reliability and validity of the FootFAST lower limb screening protocol for overuse injury.

I have read the Information Sheet concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I know that:-

1. My participation in the project is entirely voluntary;

2. I am free to withdraw from the project at any time without any disadvantage;

3. Personal identifying information in the form of my name and date of birth will be destroyed at the conclusion of the project but any raw data on which the results of the project depend will be retained in secure storage for at least five years;

4. I consent to having both pictures and video recordings made of my feet and movements and I understand that these may or may not be published in a journal as means of example, however, no images that could potentially be identifying will be included.

5. The project has been designed to minimise the risk and inconvenience to myself, the participant.

6. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve my anonymity.

7. I understand what will be required should I choose to participate.

All questions that I have are answered to my satisfaction, and I am aware that I can ask any more questions should I have them in the future.

I agree to take part in this project.

.....

.....

(Signature of participant) (Date)

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.

24 Chapter 7, Appendix 7

INFORMATION SHEET FOR

OBSERVERS

Establishing the reliability and validity of the FootFAST lower limb screening protocol for overuse injury.

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether or not to participate. If you decide to participate we thank you. If you decide not to take part there will be no disadvantage to you and we thank you for considering our request.

What is the Aim of the Project?

The aim is to identify the reliability and validity of the screening protocol: FootFAST. FootFAST is a screening protocol which assesses risk of overuse injury in the lower limb. The results of this study will be used to inform the future uses and development of the protocol.

What Type of Observers are being sought?

- If you are currently Physiotherapy clinician working in the field of musculoskeletal injury, with at least three years' experience, or,
- If you are currently a practising sports medic and have had at least three years' experience, or,
- If you are currently a NZ army medic, and have had at least three years' experience, you are eligible.

What will Observers be Asked to Do?

Should you agree to take part in this project, you will be asked to participate in four short activities. Please note that the assessments you will be making will be outlined and detailed in full. All of the pictures and videos will be given to you in one DVD.

Part 1; You will be provided with eighteen pictures of hind feet. Assess the hind foot angle using the scoring system provided.

Part 2; You will be provided with eighteen videos of single limb stance tests, Assess these performances using the scoring sheet provided.

Part 3; You will be provided with eighteen videos of a forefoot raise. Assess these performances using the scoring sheet provided.

Following your own familiarisation with the use of the protocol (approximately 20 minutes), the time commitment involved is approximately 15 minutes per participant, approximately 270 minutes in total. You will be asked to perform this same procedure on two separate occasions. This can take place wherever is convenient to you, as all that is required is a DVD player.

There is no added risk of discomfort to those who choose to participate. The inconvenience caused to you is minimal and primarily is related to the time commitment involved.

Please be aware that at any stage you may decide not to take part in the project without any disadvantage to yourself of any kind.

What Data or Information will be Collected and What Use will be Made of it?

- Personal identifying information collected for this experiment includes your name and address. This information will not be used in the analysis and will be destroyed at the completion of the project.
- The information collected will be viewed by the researchers and research student, namely, Associate Professor David McBride, Associate Professor Stephan Milosavljevic and Ms Marian Baxter. Additional members of Ms Baxter's supervisory team may also view the data to assist with the analysis.
- Data obtained from the images and video footage collected during the three activities will also be collected. The data pertains to the results of the study, and therefore will be securely stored in such a way that only those mentioned above will be able to gain access to it. Data obtained as a result of the research will be retained for **at least 5 years** in secure storage. Any personal information held on the participants [*such as name and date of birth*] may be destroyed at the completion of the research even though the data derived from the research will, in most cases, be kept for much longer or possibly indefinitely.
- The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve your anonymity.
- You have the right to view the data collected from yourself at any stage, you can contact any of the researchers using the information provided in order to do so.
- The results of the study will be made available to you upon request once data collection and analysis is completed. This is expected to be in September 2013.

Can Observers Change their Mind and Withdraw from the Project?

You may withdraw from participation in the project at any time and without any disadvantage to yourself of any kind.

What if I have any Questions?

If you have any questions about our project, either now or in the future, please feel free to contact one of the following persons:-

Marian L Baxter	Assoc Prof Stephan Milosavljevic	Assoc Professor David McBride
(03) 4762358	(03) 4797193	027 227 6546
marian.baxter@otago.ac.nz	stephan.milosavljevic@otago.ac.nz	david.mcbride@otago.ac.nz
Department of Preventive and Social Medicine University of Otago P.O. Box 913 Dunedin 9054 New Zealand	School of Physiotherapy University of Otago PO Box 56 Dunedin 9054 New Zealand	Department of Preventive and Social Medicine University of Otago P.O. Box 913 Dunedin 9054 New Zealand

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.

CONSENT FORM FOR OBSERVERS

Establishing the reliability and validity of the FootFAST lower limb screening protocol for overuse injury: A feasibility study.

I have read the Information Sheet concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I know that:-

1. My participation in the project is entirely voluntary;
2. I am free to withdraw from the project at any time without any disadvantage;
3. Personal identifying information in the form of my name and address will be destroyed at the conclusion of the project but any raw data on which the results of the project depend will be retained in secure storage for at least five years;
4. The project has been designed to minimise the risk and inconvenience to myself, the participant.
5. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve my anonymity.
6. I understand what will be required should I choose to participate.

All questions that I have are answered to my satisfaction, and I am aware that I can ask any more questions should I have them in the future.

I agree to take part in this project.

.....

.....

(Signature of participant) (Date)

This study has been approved by the University of Otago Human Ethics Committee. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479 8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.

25 Chapter 7, Appendix 8

Basic instruction for assessors

Thank you in advance for your time involved in completing this assessment.

I hope that all the instructions provided are comprehensive and simplistic, but if you are stuck or confused, please feel free to contact me:

marian.baxter@otago.ac.nz

Phone or text: 027 378 2757

Basic info:

You are provided with information from 18 participants. In each of the folders marked 1 – 18 you will find:

1. A digital image of a hindfoot
2. A video of a single leg stance performance (front and side view)
3. A video of a heel raise performance (front and side view)

You are also provided with a scoring sheet and some instructions in each folder.

The default media player of windows will likely run the videos, but if you wish to use slow motion, or cycle through frame by frame, use a media player such as VLC.

MOST IMPORTANTLY:

Please judge the performance as outlined in the information provided. There are many criteria for judging the performance of these tests, but please only stick to the criteria outlined. Always assign the worst level of scoring you observe.*

*Some of the scoring continuums are on a likert scale, and they are continuous. For example, If a participant is observed to have a very stable during the majority of their single leg stance

performance, but then has a perturbation of eversion, you may choose to place your score slightly towards ‘eversion’, but closer to ‘stable’.

When you are finished filling in all the scoring sheets from all three folders, please return them via email. The scoring sheets are the only documents I need returned.

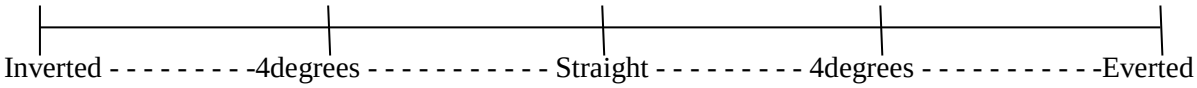
NB: You can work through Tests 1 , 2 and 3 in any order you like and take as much time as you feel necessary. However, please complete each test in numerical order as indicated in the files.

Recording Sheet:

Assessment of Hind foot position

Observe the hindfoot shown in the pictures, numbered one to ten. Indicate on the scale shown where you think the position of the hindfoot is of the participants' **FOOT AS INDICATED**. You can place your mark absolutely anyway along the line.

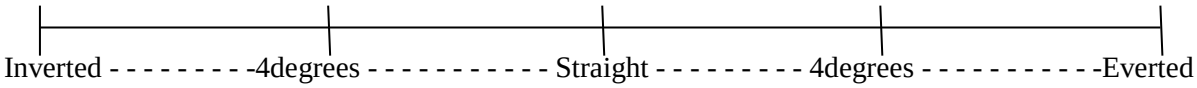
Picture 1: **RIGHT**



Inverted - - - - - - - - -4degrees - - - - - - - - - - - Straight - - - - - - - - - 4degrees - - - - - - - - - - -Everted

>4 degrees Inversion Eversion >4 degrees

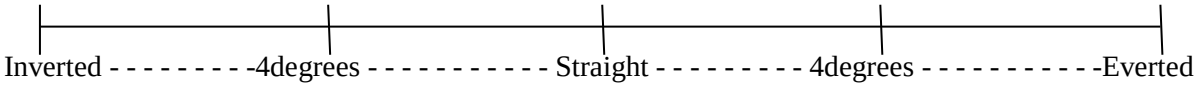
Picture 2: **RIGHT**



Inverted - - - - - - - - -4degrees - - - - - - - - - - - Straight - - - - - - - - - 4degrees - - - - - - - - - - -Everted

>4 degrees Inversion Eversion >4 degrees

Picture 3: **LEFT**



Inverted - - - - - - - - -4degrees - - - - - - - - - - - Straight - - - - - - - - - 4degrees - - - - - - - - - - -Everted

>4 degrees Inversion Eversion >4 degrees

****And so on for all 18 participants.**

Recording Sheet:

Assessment of Single leg stance

Observe the video of a single leg stance performance, numbered one to ten. Indicate on the scale shown how good you think the performance is. The definition of each classification is given:

Stable: The foot remains balanced, without perturbation or oscillations

Mildly unstable: On occasion, the foot tends to oscillate or correct itself

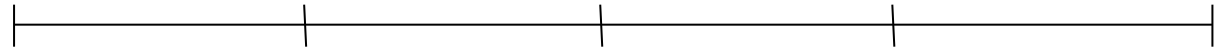
Unstable: The foot is continuously oscillating and there may be a reduction in the ground-surface contact at times

Laterally unstable: The foot continuously falls towards the lateral edge of the foot, losing ground contact on the medial aspect.

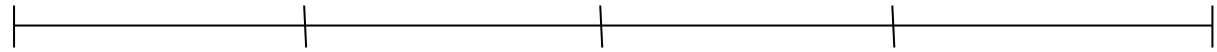
Cannot perform task: As indicated, if the participant is not balanced for at least 5 seconds.

As noted in the 'Basic Instructions' document, always assign the worst score for the performance you observe. You can place your mark absolutely anyway along the line.

Video 1:



Video 2:



****And so on for all 18 participants.**

Recording Sheet:

Assessment of Heel Raise

Observe the video of a heel raise performance, numbered one to ten. Indicate on the scale shown how good you think the performance is. **The performance ends when the participant has reached their maximum height (i.e. transition through heel raise, not maintenance of).** The definition of each classification is given:

Smooth: The participant can move in a controlled manner into a high position on their toes.

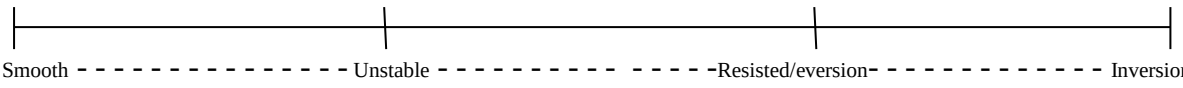
Unstable: The foot may or may not be able to reach a good height, but the motion is not smooth.

Resisted/ eversion: The foot struggles moving upwards, may even tend to fluctuate up and down, and there may be pronounced "falling inwards" where the foot twists in an inwards manner.

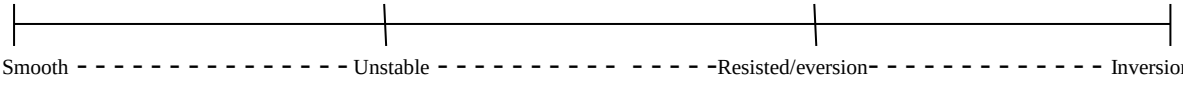
Inversion: The foot repeatedly "falls outwards" where the foot loses contact with the inside (medial) edge of ground surface, **and** height cannot be reached.

As noted in the 'Basic Instructions' document, always assign the worst score for the performance you observe. You can place your mark absolutely anyway along the line.

Video 1:



Video 2:



****And so on for all 18 participants.**

26 Appendix 9, Ethical approval

22 May 2012

Academic Services

Manager, Academic Committees, Mr Gary Witte

12/120

Assoc. Prof. S Milosavljevic

School of Physiotherapy

Dear Assoc. Prof. Milosavljevic,

I am writing to let you know that, at its recent meeting, the Ethics Committee considered your proposal entitled "**Establishing the reliability and validity of the FootFAST lower limb screening protocol for overuse injury: A feasibility study.**".

As a result of that consideration, the current status of your proposal is:- **Approved**

For your future reference, the Ethics Committee's reference code for this project is:- **12/120**.

The comments and views expressed by the Ethics Committee concerning your proposal are as follows:-

While approving the application, the Committee would be grateful if you would respond to the following:

Approval is for up to three years from the date of this letter. If this project has not been completed within three years from the date of this letter, re-approval must be requested. If the nature, consent, location, procedures or personnel of your approved application change, please advise me in writing.

20 August 2012

Academic Services

Manager, Academic Committees, Mr Gary Witte

12/216

Assoc. Prof. D McBride

Department of Preventive and Social Medicine

Dunedin School of Medicine

Dear Assoc. Prof. McBride,

I am writing to let you know that, at its recent meeting, the Ethics Committee considered your proposal entitled "**Establishing the reliability and validity of the FootFAST lower limb screening protocol for overuse injury**".

As a result of that consideration, the current status of your proposal is:- **Approved**

For your future reference, the Ethics Committee's reference code for this project is:- **12/216**.

The comments and views expressed by the Ethics Committee concerning your proposal are as follows:-

While approving the application, the Committee would be grateful if you would respond to the following:

Approval is for up to three years from the date of this letter. If this project has not been completed within three years from the date of this letter, re-approval must be requested. If the nature, consent, location, procedures or personnel of your approved application change, please advise me in writing.