A mixed methods study to explore the diagnostic accuracy and acceptability of the tuning fork test in the detection of ankle fractures.

Anne Welling

The thesis is submitted in partial fulfilment of the requirements for the award of the degree of Professional Doctorate in Nursing of the University of Portsmouth

January 2012
Abstract

Aim and methodology:

Ankle injuries account for 8% of all minor injuries attending emergency departments in the United Kingdom and the Ottawa ankle rules were introduced to assess the need for x-ray in the early 1990s (Stiell et al 1992). Although the rules are said to have reduced the number of ankle x-rays requested the frequency of fractures in the population still receiving x-rays is only 15% nationally. This study aims to assess whether the tuning fork can increase the diagnostic accuracy of the Ottawa ankle rules when used on twisting ankle injuries by multiple operators in multiple emergency care settings.

A mixed methods study conducted in two phases was undertaken. Phase one consisted of a diagnostic test study using the Ottawa ankle rules in conjunction with the tuning fork test on patients already screened as being Ottawa positive to the 'malleolar' zone and requiring an x-ray of their ankle. Patients aged 12 years or over who had sustained an ankle injury by a twisting mechanism were eligible to take part. Patient age, gender, ethnicity, and previous history of injury or presence of distracting injuries, degree of swelling, and role of operator were all considered potential variables for an accurate tuning fork test, and these were analyzed individually and in a multiple logistical regression model to assess for predictor variables of a correct tuning fork test.

Phase two of the study included a series of focus group discussions to explore participant and clinician experiences of the tuning fork test. Data was analyzed using thematic analysis.

Results

Data was collected for 2-years and 1313 patients were included in the final analysis. 56% of the study participants were male. Mean age was 34 years (range 12-91). 98% were of white ethnic origin. 210 (16%) were diagnosed with fractures, of which 38 were deemed to be not clinically significant. The tuning fork had a diagnostic accuracy of 56% (95% CI 53-58), NPV 96% (95% CI 94-97), sensitivity 84% (95% CI 78-89) and specificity 51% (95% CI 48-54). X-rays could have been reduced by 47% but this was at the expense of missing 29 ‘clinically significant’ fractures. However, seven of these were managed as soft tissue injuries and in nine the initial assessment of tenderness did not match
the site of the fracture. A total of 113 clinicians (nurses & doctors) were involved in performing the tuning fork test independently. Patient age (adjusted OR 1.021, p. <0.001) and role of the operator (adjusted OR 1.595, p. 0.003 for nurse) were the only predictors of an accurate test.

Ten patients and ten clinicians attended the focus group discussions in phase two of the study. Patients and clinicians appeared to accept the tuning fork as a method for assessment provided adequate explanation was given. Patients claimed the tuning fork test was not painful but had a similar sensation to that of a ‘Tens’ machine. There were differences in opinion between the two groups as to whether the tuning fork was accurate or not and clinicians held the perception that patients expect an x-ray when they present with an ankle injury, whereas patients disagreed with this. Patients were fully aware of the dangers of x-rays and stated that a reduction in x-rays was one of the main potential benefits of the study.

Conclusions

This is the largest study to investigate the accuracy of the tuning fork to detect fractures, not only in the size of the study population but the number of clinicians involved. It is also the first to report inconclusive Ottawa ankle rule and tuning fork test results. It is unlikely that the lower sensitivity will be accepted by patients and clinicians. Further research to assess inter-operator reliability is recommended before implementing the tuning fork test into clinical practice.
Contents list

Declaration.................................................................................................................. vi

List of tables................................................................................................................. vii

List of figures................................................................................................................ ix

Abbreviations.............................................................................................................. xi

Acknowledgements & Dedication............................................................................. x

Dissemination............................................................................................................. xi

Chapter 1 - INTRODUCTION...................................................................................... 1

Chapter 2 - BACKGROUND & RATIONALE FOR THE STUDY ......................... 6
  2.1: Aetiology of ankle injuries .............................................................................. 6
  2.2: Ankle anatomy & ankle sprains................................................................. 7
  2.3: Ankle fractures............................................................................................ 15
  2.4: Ankle fractures in children.......................................................................... 17
  2.5: Examination of the ankle........................................................................... 19
  2.6: Ankle X-rays.................................................................................................. 21
  2.7: Development of the Ottawa ankle rules................................................... 22
  2.8: Implementation of the Ottawa ankle rules............................................ 27
  2.9: Operator experience .................................................................................. 30
  2.10: Reduction in x-rays.................................................................................. 31
  2.11: Effect of the Ottawa ankle rules on waiting times............................. 32
  2.12: Alternative ankle rules............................................................................. 33
  2.13: Ottawa ankle rules & the tuning fork..................................................... 41

Chapter 3 – A SYSTEMATIC REVIEW OF THE LITERATURE ON THE USE OF THE TUNING FORK TO DETECT FRACTURES .............................................. 45
  3.1: History of tuning fork.................................................................................. 45
  3.2: Review method.............................................................................................. 47
3.3: Research question & hypothesis................................................................. 62

Chapter 4 – PHASE ONE: METHODOLOGY FOR ASSESSING THE
DIAGNOSTIC ACCURACY OF THE TUNING FORK TEST IN THE
DETECTION OF FRACTURES TO THE ANKLE................................................. 70
  4.1: Ethical & Regulatory Approval ................................................................. 70
  4.2: Participants ............................................................................................... 74
  4.3: Index & reference test............................................................................... 80
  4.4: Training ...................................................................................................... 82
  4.5: Statistical methods ................................................................................... 83
  4.6: Study website ........................................................................................... 89
  4.7: Findings .................................................................................................... 90
  4.8: Discussion of findings ............................................................................... 109
  4.9: Strengths & weaknesses of this phase of the study................................. 119

Chapter 5 – PHASE TWO: EXPLORATION OF PATIENT AND CLINICIAN
EXPERIENCES OF THE TUNING FORK TEST VIA FOCUS GROUP
DISCUSSIONS .................................................................................................... 127
  5.1: Patient focus group discussions............................................................... 127
  5.2: Clinician focus group discussions ............................................................ 164
  5.3: Discussion of findings .............................................................................. 188
  5.4: Strengths & weaknesses of this phase of the study................................. 190
  5.5: Conclusion ............................................................................................... 195

Chapter 6 - DISCUSSION ................................................................................... 197
  6.1: Overview of study methodology ............................................................... 197
  6.2: Clinician experience ............................................................................... 200
  6.3: Population & predictor variables ............................................................... 201
  6.4: Inconclusive results ................................................................................. 203
  6.5: Accuracy ................................................................................................... 204
  6.6: Missed fractures ...................................................................................... 206
  6.7: Potential benefits ..................................................................................... 207
  6.8: Limitations of study ................................................................................. 207

Chapter 7 - CONCLUSION & IMPLICATIONS FOR PRACTICE ..................... 209
  7.1: Conclusion ................................................................................................ 209
7.2: Implications to practice & recommendations for future research……218

Chapter 8 – REFLECTION......................................................................................... 220
  8.1: On undertaking Professional Doctorate training .................................. 220
  8.2: On my role as a practitioner-researcher............................................... 225
  8.3: On the future.......................................................................................... 226

Chapter 9 - REFERENCES...................................................................................... 228

Chapter 10 - APPENDICES.................................................................................. 245
Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

Word count: 43968
List of tables

TABLE 1: PREDICTOR VARIABLES OF A FRACTURE ACCORDING TO THE LEIDEN & UTRrecht ANkle RULES ................................................................. 36


TABLE 3: DATABASES SEARCHED - DESCRIPTION & DATES ................................ 48

TABLE 4: SUMMARY OF DIAGNOSTIC RESEARCH PAPERS USING AUSCULTATORY PERCUSSION OR TUNING FORK TO DETECT FRACTURES ........................................... 51

TABLE 5: MAIN FINDINGS, SENSITIVITY, SPECIFICITY, STRENGTHS & WEAKNESS OF DIAGNOSTIC RESEARCH PAPERS ................................................................. 52

TABLE 6: SUMMARY OF CASE STUDIES, LITERATURE REVIEW & NARRATIVES ON AUSCULTATORY PERCUSSION & TUNING FORKS IN THE DETECTION OF FRACTURES ............................................................................. 53

TABLE 7: RESULTS OF WIDER ET AL (2008) COMPARING DIFFERENT TUNING FORKS IN THE DETECTION OF STRESS FRACTURES ........................................................................ 61

TABLE 8: FACTORS AFFECTING MEASURES OF DIAGNOSTIC ACCURACY .............. 67

TABLE 9: STARD CHECKLIST FOR REPORTING OF STUDIES OF DIAGNOSTIC ACCURACY .......................................................................................................................... 69

TABLE 10: ETHICAL PRINCIPLES RELEVANT TO THE STUDY .................................. 72

TABLE 11: INCLUSION & EXCLUSION CRITERIA ....................................................... 78

TABLE 12: DATA COLLECTED BY THE ASSESSING CLINICIAN .................................. 84

TABLE 13: 3x3 TABLE ............................................................................................... 87

TABLE 14: DEMOGRAPHIC PROFILE & PRESENTATION CHARACTERISTICS OF THE STUDY PARTICIPANTS .................................................................................................... 96

TABLE 15: DISTRIBUTION OF FRACTURE LOCATION ............................................ 98

TABLE 16: DIAGNOSIS OUTCOME FOR FRACTURES: TUNING FORK VERSUS X-RAY .... 98

TABLE 17: 2x2 TABLE SHOWING RELATIONSHIP BETWEEN TUNING FORK TEST & X-RAY ................................................................................................................. 99

TABLE 18: SUMMARY OF THE 27 CLINICALLY SIGNIFICANT FRACTURES MISSED BY THE TUNING FORK ........................................................................................................ 101

TABLE 19: VARIABLES PREDICTIVE OF AN ACCURATE TUNING FORK TEST .......... 103

TABLE 20: ACCURACY OF THE TUNING FORK BY CLINICIAN ROLE: NURSES V DOCTORS ..................................................................................................................... 104

TABLE 21: ACCURACY OF THE TUNING FORK IN THE PRESENCE OF SWELLING .... 105

TABLE 22: FACTORS PREDICTIVE OF A "POSITIVE" TUNING FORK TEST ............ 107

TABLE 23: RESULT OF A "POSITIVE" TUNING FORK TEST BY CLINICIAN ROLE ........ 108

TABLE 24: DATES & TIMES OF THE PATIENT FOCUS GROUPS .................................. 134
List of tables

TABLE 26: PATIENT CHARACTERISTICS ATTENDING THE FOCUS GROUPS & POSTING COMMENTS ON THE WEBSITE FORUM ................................................................. 144
TABLE 27: PATIENT & CLINICIAN INTERPRETATION OF THE TUNING FORK TEST & THE X-RAY RESULT ........................................................................ 145
TABLE 28: INITIAL THEMES FROM PATIENT FOCUS GROUPS ...................... 146
TABLE 29: KEY THEMES .......................................................................... 147
TABLE 30: INITIAL THEMES - CLINICIAN FOCUS GROUPS ..................... 170
TABLE 31: KEY THEMES .......................................................................... 171
List of figures

FIGURE 1: DIAGRAMMATIC REPRESENTATION OF THE BONES OF THE ANKLE .......... 8
FIGURE 2: SURFACE ANATOMY OF THE ANTERIOR ANKLE .................................. 8
FIGURE 3: DIAGRAMMATIC REPRESENTATION OF DORSI & PLANTAR FLEXION .......... 9
FIGURE 4: DIAGRAMMATIC REPRESENTATION OF INVERSION & EVERSION .......... 10
FIGURE 5: DIAGRAMMATIC REPRESENTATION OF THE MEDIAL COLLATERAL LIGAMENT 11
FIGURE 6: DIAGRAMMATIC REPRESENTATION OF THE LATERAL COLLATERAL LIGAMENT ................................................................. 12
FIGURE 7: WEBER CLASSIFICATION OF ANKLE FRACTURES ................................. 15
FIGURE 8: SMALL AVULSION FRACTURE TO THE LATERAL MALLEOLUS .................. 16
FIGURE 9: SALTER-HARRIS FRACTURE CLASSIFICATION ................................... 18
FIGURE 10: PLAIN AP & LATERAL ANKLE X-RAYS ............................................. 21
FIGURE 11: OTTAWA ANKLE RULES .................................................................... 24
FIGURE 12: MNEMONIC 44-55-66-PM TO AID RETENTION OF THE OARs ............ 28
FIGURE 13: FLAT END TUNING FORK ................................................................... 41
FIGURE 14: SUMMARY OF SEARCH RESULTS ..................................................... 50
FIGURE 15: SCREEN SHOT OF TUNING FORK STUDY HOME PAGE ..................... 89
FIGURE 16: SUMMARY OF ANKLE INJURIES ATTENDING STUDY SITES A & B DURING STUDY PERIOD ............................................................................. 91
FIGURE 17: SUMMARY OF PATIENT RECRUITED & REMOVED FROM THE STUDY .... 92
FIGURE 18: RECRUITMENT TOTAL BY MONTH .................................................. 93
FIGURE 19: RECRUITMENT BY STUDY .................................................................. 94
FIGURE 20: AGE RANGE OF PARTICIPANTS ....................................................... 95
FIGURE 21: SCREEN SHOT OF STUDY FORUM WEBPAGE .................................. 133
FIGURE 22: SUMMARY OF RESPONSE RATE FOR THE FIRST FOUR FOCUS GROUPS .. 141
FIGURE 23: SUMMARY OF RESPONSE RATE FOR 5TH FOCUS GROUP ............... 143
Abbreviations

AMED – Allied & Complimentary Medicine Database
AP – Antero-posterior
BAEM – British Association of Emergency Medicine
BUPA – British United Provident Association
CCF – Central Commissioning Facility
CI – Confidence Interval
CINAHL – Cumulative Index to Nursing & Allied Health Literature
CLRN – Comprehensive Local Research Network
CT – Computerised tomography
DH - Department of Health
ED – Emergency Department
ENP – Emergency Nurse Practitioner
GP – General Practitioner
IQ – Interquartile Range
IRAS – Integrated Research Application System
MSc – Master of Science degree
MRI – Magnetic Resonance Imaging
NHS – National Health Service
NIHR – National Institute of Healthcare Research
NPV – Negative predictive value
OARs – Ottawa ankle rules
OR – Odds Ratio
PACs – Picture Archiving Communications System
PhD – Doctor of Philosophy
RfPB – Research for Patient Benefit
STARD – Standards for reporting diagnostic accuracy
UK – United Kingdom
USA – United States of America
Acknowledgements & Dedication

I would like to thank NHS South Central (formerly the Wessex Deanery) for supporting me to undertake this Professional Doctorate whilst on the consultant practitioner training programme and after securing my role as a consultant nurse. I would like to thank all the clinicians involved in obtaining consent and performing the tuning fork test, as without them this study would not have been possible.

A big thank you goes to Bernie Higgins without whose patience the study data may never have been analysed! I would like to thank my academic supervisors Dr Ann Dewey and Dr Sally Kilburn for their guidance, feedback and critical appraisal throughout the program of study, from you both I have learnt a great deal.

Finally I wish to thank my husband George who without any academic or medical knowledge has given me the support I needed to complete this study. He has provided the foundation on which I have been able to rely not only for this period of study but throughout my whole career. Without his love, support, understanding and words of encouragement none of this would have been possible, and for this I will be forever grateful.

I dedicate this thesis to my husband George.
Dissemination


**Welling, A.** (2011). *The Tuning fork test: Is it a simple adjunct to the Ottawa ankle rules?* Workshop 2.2, at the RCN Emergency Care Association annual conference and exhibition, November 11, 2011
Chapter 1 - INTRODUCTION

There is no data available on the precise number of ankle fractures occurring in the United Kingdom (UK) but it is estimated that they account for 3-5% of all attendances to emergency departments (ED) in the UK (Trundle 1997) with the incidence increasing steadily over the past twenty years (Koehler & Eiff 2010). The most common ankle injury is a sprain, which is a tearing of the ankle ligaments and the British United Provident Association [BUPA] Health Information Team (BUPA 2010) estimate that 1.5 million ED attendances are for ankle sprains.

Following an ankle injury the aim of clinical examination is to differentiate between a sprain and a fracture. This is achieved by obtaining a thorough history of the injury from the patient and undertaking an assessment to identify potential sites of bony tenderness. If bony tenderness is identified an x-ray is requested to confirm or exclude the presence of a fracture. Research in the 1980s revealed that the majority of patients presenting to EDs with ankle injuries received an x-ray, even though 85-90% were later diagnosed with a sprain (Bachmann, Kolb, Toller, Steurer & Riet 2003). In response to this Stiell et al (1992) undertook a series of clinical trials to identify key variables predictive of an ankle fracture. Key predictor variables were identified as pain to the ankle zone and tenderness to either the posterior aspect of the lateral or medial malleolus, the distal 6cm of the fibula shaft or the inability to walk immediately after the injury or in the ED. These key predictor variables later became known as the Ottawa ankle rules (OARs) and have since become one
of the most validated clinical decision tools used in practice. A meta-analysis by Bachmann et al (2003) identified that the pooled sensitivity of 27 studies using the OARs was 98%, with 95% confidence interval (CI) of 96-99%, and Bachmann et al (2003) confirmed that only 2% of patients confirmed as OARs negative had a fracture. However, this high sensitivity is at the expense of a low specificity with Bachmann et al (2003) reporting a median specificity of 39% (interquartile range 28-48) across the 27 studies included in their meta-analysis.

The low specificity of the OARs has also been noted locally. An audit, undertaken in 2006, at an ED on the south coast of the UK included fifty consecutive patients presenting with an ankle injury, 42 were sent for x-ray. Only six fractures were diagnosed which represents a specificity of only 18%. This suggests that at both a local and national level there is the potential to improve the specificity of the OARs, but only if the high sensitivity can be maintained.

A recent study suggests that a tuning fork might increase the specificity of the OARs whilst maintaining the high sensitivity. This study by Dissmann and Han (2006) used a small sample size (N=49), reported wide CI for sensitivity, only included patients with lateral malleolar tenderness without swelling, and the tuning fork test was performed by a single operator, which limits the generalizability of the results of this study. Nevertheless, the findings are interesting and provide credible results to further investigate the potential use of tuning forks to improve the diagnosis of ankle fractures.
A search of the literature revealed a limited evidence base to support the claims from the Dissmann and Han (2006) study so it was therefore decided to undertake a larger well conducted clinical study to address the limitations of the Dissmann and Han (2006) study. In order to ensure replication of the process in the clinical setting the study would need to include patients with tenderness and swelling to either the lateral or the medial malleolus, and involve multiple clinicians. If proven to increase the specificity of the OARs the tuning fork test could provide a number of future benefits to the National Health Service [NHS], including reducing the number of ankle x-rays requested, and therefore reducing costs, whilst benefits for patients could include reducing radiation exposure and waiting times.

This thesis, which forms part of the Professional Doctorate in Nursing, describes a two-phase study that used a mixed methods approach to include both quantitative and qualitative data collection techniques. It comprises a quantitative diagnostic test to assess whether a tuning fork can be used to increase the diagnostic accuracy of the OARs, and a qualitative enquiry in the form of focus group discussions to explore the experiences and views of patients and clinician on receiving and using the tuning fork test in clinical practice. This thesis contains eight further chapters and a summary of each is given below.

- Chapter 2 summarises the background and rationale of the study. It includes clinical examination of the ankle following injury and outlines the relevant anatomy in relation to the mechanism of injury. It also describes
the common injuries that occur from various mechanisms of injury. A detailed discussion of the development of the OARs is also included.

- Chapter three contains a brief outline of the history of the tuning fork followed by a systematic search and critical appraisal of the literature which sought to establish the strength of evidence base to support the use of the tuning fork to diagnose fractures in clinical practice.

- Chapter four describes phase one of the study and includes the methodology and results of a large diagnostic test study which set out to assess the accuracy of the tuning fork test in diagnosing fractures to the ankle.

- Chapter five describes phase two of the study and includes the methodology and findings of patient and clinician focus group discussions which set out to explore patient experiences of receiving the tuning fork test and clinician experiences of administering the tuning fork test.

- Chapter six contains a discussion of the whole study where the contribution of the quantitative and qualitative data is examined together with reference to the current evidence base. The limitations of the study are also discussed.

- Chapter seven contains the study conclusion, and outlines what this study contributes to current evidence of knowledge. The implications for practice are also discussed.

- Chapter eight contains a reflective account of the skills and knowledge gained during the Professional Doctorate training and outlines recommendations for future research.
- Chapters nine, ten and eleven contain the references, appendices and copies of publications respectively.
Chapter 2 - BACKGROUND & RATIONALE FOR THE STUDY

In order to understand the rationale for the study, it is necessary to describe the process of assessing an ankle following injury, the anatomy involved, and how common injuries occur. This chapter will outline the aetiology of ankle injuries in the UK, outline the process of clinical examination, with reference to relevant anatomy, and describe the common ankle injuries that can occur following simple twisting mechanisms. Although the Ottawa ankle rules (OARs) are one of the most validated clinical decision rules used in clinical practice this section will also include a description of how they were developed and outline how they compare to other ankle rules used in clinical practice.

2.1: Aetiology of ankle injuries

Despite an in-depth review of internet sites that focus on health statistics there appears to be no data on the number of ankle injuries that occur in the UK, but they are said to account for 3-5% of all attendances to emergency departments (ED) in the UK (Trundle 1997). The most common mechanism of injury to the ankle occurs when the ankle twists, usually resulting in a sprain or a fracture. Ankle sprains are more common than fractures and according to BUPA (2010) they account for 1.5 million ED attendances annually in the UK alone. This equates to a total of 5600 people reportedly spraining their ankle every day in the UK (Pijenberg et al 2000), and accounts for a quarter of all sports injuries (Struijs & Kerkhoffs 2010).
In a systematic review of 227 epidemiology studies on sports injuries, which included papers from eight different countries, Fong, Hong, Chan, Yung and Chan (2007) identified that the ankle was second only to the knee as the body part most injured during sporting activities and that the ankle sprain was the most common injury in 33 out of 43 different sports reviewed. There are a number of common mechanisms of injury for an ankle injury and when Eggli, Sclabas, Eggli, Zimmermann, and Exadaktylos (2005) reviewed patients presenting with simple twisting ankle injuries they found that 52% (184/354) injured their ankle during sport, 21% (74/354) injured their ankle whilst at work, 13% (46/354) whilst at home, and 8% (28/354) occurred in road traffic accidents.

2.2: Ankle anatomy & ankle sprains

The ankle is made up of three bones – the tibia, the fibula and the talus (Mai & Cooper 2010). The tibia is the larger of the two bones in the lower leg and forms the inside or medial part of the ankle. Figure 1 shows the position of the bones of the ankle. Distally the tibia forms the bony prominence on the inside of the ankle, known as the medial malleolus (Bickley 2003). The fibula is the smaller of the two shin bones, and lies on the outside or lateral part of the shin. Distally the fibula forms the lateral bony prominence of the ankle joint known as the lateral malleolus (Mai & Cooper 2010).
Distal to the fibula and tibia is the talus, which in turn articulates with the calcaneum, commonly known as the heel bone (Mai & Cooper 2010). Figure 2 shows the surface anatomy of the ankle.
The three bones of the ankle form three separate joints - the talocrural, the tibio-fibula and the subtalar, which allow the ankle to articulate through four planes of movement, namely eversion, inversion, dorsi-flexion and plantar flexion (Sports Injury Clinic 2010). The talocrural joint is a hinge joint formed by the distal ends of the fibula and tibia where they articulate with the upper surface of the talus, allowing the ankle to dorsi and plantar-flex. Figure 3 shows a diagrammatic representation of dorsi and plantar flexion.

The tibio-fibula joint lies between the lower surface of the tibia and fibula, and the subtalar joint includes the articulating surface of the talus and the calcaneum. The subtalar joint allows inversion and eversion of the ankle. Figure 4 shows the movements of inversion and eversion.
A forced inversion or eversion through any mechanism will result in injury, the severity of which will depend on the forces exerted on the ligaments and bones.

The main ligaments injured following simple twisting mechanisms are the lateral collateral and the medial collateral ligaments (Evans & Schucany 2006). The medial collateral ligament lies on the medial aspect of the ankle and is also commonly known as the deltoid ligament (Bickley 2003). It is thicker and stronger than the lateral collateral ligament, which lies on the lateral aspect of the ankle. The medial collateral ligament fans out to cover the distal end of the tibia and inner aspect of the talus, navicular and calcaneum (Bickley 2003). The deep tibiotalar band of the medial collateral ligament is the most important for ankle stability (Barrie & Lishman 2010). Figure 5 shows a diagrammatic representation of the medial collateral ligament.
Isolated medial collateral ligament sprains are rare but when they do occur they are usually associated with significant joint instability or fracture to the medial malleolus (Sports Injury Clinic 2010).

The lateral collateral ligament lies on the lateral aspect of the ankle and is the name given to a ligament that has three distinct bands - all of which attach to the lateral malleolus. The anterior tibia-fibular ligament passes from the lateral malleolus to the front of the talus; the calcaneo-fibular ligament connects the calcaneum to the lateral malleolus (Sport Injury Clinic 2010); and the posterior talofibular ligament attaches from the back of the lateral malleolus to the rear of the talus. The role of the lateral collateral ligament is to prevent excessive...
inversion. Figure 6 shows a diagrammatic representation of the three bands of the lateral collateral ligament.

![Diagram of the lateral collateral ligament]

Key: ATFL = Anterior tibio-fibular; PTFL = posterior tibio-fibula; CFL = Calcaneo-fibular.

Figure 6: Diagrammatic representation of the lateral collateral ligament

Retrieved April 29, 2011 from [http://www.hawaii.edu/medicine/pediatrics/pemxray/v3c03b.jpg](http://www.hawaii.edu/medicine/pediatrics/pemxray/v3c03b.jpg)

A sprain to the lateral collateral ligament is the most common ligamentous injury that occurs to the ankle, firstly because it is much weaker than the larger medial collateral ligament and secondly as the medial malleolus is shorter in length than the lateral malleolus the ankle has a tendency to invert rather than evert (Garrick 1997).

In addition the ankle contains two other important ligaments that can be injured following twisting mechanisms, namely the syndesmosis and the interosseous
tibio-fibular ligament (Evans & Schucany 2006). The syndesmosis is a strong ligament that joins the distal portion of the tibia and fibula together and the interosseus tibiofibular ligament lies between the shafts of the tibia and fibula. The position of the interosseus tibiofibular ligament can be seen in Figure 6. When the syndesmosis or interosseus tibiofibular is damaged it is often referred to as a high ankle sprain, and is associated with widening of the space between the tibia and fibula and results in an unstable ankle, interfering with dorsiflexion (Evans & Schucany 2006).

Ankle sprains can be divided into three different grades dependent on the severity of the injury and distinguished by the amount of swelling, bruising, joint stiffness and ability to walk on the affected limb (Young & Ho 2011). Toung and Ho (2011) define the different grades of ankle sprain as

- **Grade 1** - occur as a result of some stretching or minor tearing of the lateral ankle ligaments only. They are associated with only minor swelling around the lateral malleolus and a degree of difficulty walking due to pain and joint stiffness. Management of these injuries is limited to rest and elevation for the first 24 to 48 hours

- **Grade 2** - occur as a result of moderate tearing of the ligament fibres and cause moderate to severe pain on walking as well as bruising, swelling and stiffness to the joint.

- **Grade 3** - occur when there is total rupture of a ligament causing gross instability of the ankle joint. This will be associated with gross swelling,
severe pain and extensive bruising. A grade three sprain can result in dislocation of the ankle joint.

The aim of treatment following an ankle sprain is to return to full mobility as quickly as possible. This generally involves the use of rest, ice, elevation and movement of the ankle to ensure the scar tissue, which forms after a ligament becomes injured, is not allowed to become tight and shortened (Cluett 2010). Whether to support a sprained ankle with a bandage or splint is open to debate as this has been known to lead to a stiff ankle, delay mobility and make the ankle more prone to further injury (Cluett 2010). In a Cochrane systematic review of the management of ankle injuries Kerkhoffs et al (2002) reviewed nine studies and included 892 patients. Kerkhoffs et al (2002) reviewed the benefits of elastic bandage, lace-up and rigid ankle support and taping on patients with ankle sprains. They found that a lace-up ankle support was better for swelling in the short-term, but that a rigid ankle support resulted in a shorter time to return to work. In addition they identified that an elastic bandage resulted in fewer complications but was associated with more instability of the ankle, and a slower return to work and sport when compared to a rigid ankle support. Kerkhoffs et al (2002) state that due to the variety of treatments available and inconsistency in follow-up of the patients they were unable to make a definitive conclusion as which was the most effective treatment clinically and for cost efficiency.

Locally grade 1 and 2 ankle sprains are managed with rest, ice and elevation plus analgesia and crutches for patients who are unable to bear weight in the
ED. Grade 3 sprains are often managed in a below knee plaster cast or air cast boot for one week and then followed up by either the orthopaedic team or physiotherapy.

2.3: Ankle fractures

In addition to sprains any of the bones in the ankle can fracture. The lateral malleolus is the most common malleoli injured and fractures to this site have been classified to indicate severity. Locally ankle fractures are classified under the Weber classification (Gaillard 2008). Figure 7 shows a diagrammatic representation of the Weber classification.

![Figure 7: Weber classification of ankle fractures](http://4.bp.blogspot.com/_as7Ap63dYXM/S2qGPk7HmhI/AAAAAAAABF4/3f-n-j4JJxE/s320/danis_webber_classification_ankle_fractures.png)

‘Weber A’ defines a fracture that occurs below the tibio-talar joint line. Ligaments will remain intact but there may be an associated medial malleolus
fracture. Although this is a stable fracture larger fracture segments may require internal fixation (Gaillard 2008). The Weber A classification also includes small avulsion fractures which occur when the ligament is stretched and pulls a small piece of bone away. Small avulsion fractures, measuring less than 3mm in depth, to the lateral malleoli can be considered not to be clinically significant, as they do not require immobilisation and can be safely managed as sprains (Stiell et al 1992). Figure 8 shows an ankle x-ray with the arrow indicating a small avulsion fracture.

Figure 8: Small avulsion fracture to the lateral malleolus

‘Weber B’ defines a fracture occurring at the level of the tibio-talar joint line where the fracture extends laterally and superiorly up the fibula. The syndesmosis may be intact or partially torn, but there is no widening of the tibio-fibula joint. There may be associated tenderness to the medial aspect of the ankle and a fracture to the medial malleolus (Gaillard 2008).

‘Weber C’ defines a fracture above the tibio-talar joint line. The syndesmosis will be disrupted and widened, and there will be an associated injury to the deltoid ligament or fracture to the medial malleolus.

2.4: Ankle fractures in children

Specific to children fractures can also involve the growth plates of the tibia and fibula. These fractures are grouped under the Salter-Harris classification and include five distinct classifications (Cluett 2010).

- **Salter-Harris Type I** - occur predominantly in younger children with the fracture going directly through the growth plate. X-rays will appear normal and complications are rare.
- **Salter-Harris Type II** – The fracture line starts across the growth plate then continues up the shaft of the bone and away from the joint. These fractures mainly occur in older children and complications are rare.
- **Salter-Harris Type III** – The fracture line starts through the growth plate and then exits the end of the bone through the joint. As a result the
joint cartilage is damaged. These fractures tend to occur in older children.

- Salter-Harris Type IV – The fracture starts above the growth plate, before crossing it and exiting into the joint. Again the joint is disrupted and these fractures often require surgery to stabilise.
- Salter-Harris Type V – In this fracture the growth plate is crushed. This type of fracture can result in poor bone alignment and affect the future growth of the bone.

Figure 9 shows a diagrammatic representation of the Salter-Harris fracture classification.

![Salter-Harris Classification of Growth Plate Injuries](http://www.niams.nih.gov/Health_Info/Growth_Plate_Injuries/graphics/growth-plate.jpg)

In addition, children can also be diagnosed with ‘incomplete’ fractures, and these are commonly known as greenstick, torus and buckle fractures (Long n.d.). Each of these occur in the shaft of the bone and can affect the fibula or the tibia. In each the bone has a tendency to bend and the cortex of the bone
buckle or break on one side. These fractures result in pain and swelling and require a short period of immobilisation but generally cause no long term complications (Long n.d.).

In order to assess the severity of any ankle injury the clinician needs to undertake a systematic clinical examination of the ankle and the next sub-section will discuss the standard examination required when a patient presents with an ankle injury.

**2.5: Examination of the ankle**

Following injury to the ankle it is important for the clinician to assess the extent of the injury and this is achieved firstly by obtaining a thorough history of the mechanism of the injury from the patient before progressing with clinical examination. The history of the injury will inform the clinician when the injury occurred, the symptoms experienced and any first aid measures carried out by the patient. In addition the clinician can identify the onset, location, duration, and character of any pain, as well as any associated, aggravating or relieving factors the patient may have identified. Specific to an ankle injury the majority of patients will describe a forced eversion or inversion of the ankle or foot. Examples from personal practice include twisting the ankle going down a kerb, walking on uneven pavements, or falling whilst in high-heel shoes.
Clinicians commonly use the systematic process of ‘look, feel, move’ to examine traumatic injuries to the limbs (Bickley 2003). ‘Looking’ at the ankle and comparing to the non-injured side can offer a clue as to the extent of the injury by the degree of swelling, bruising, presence of wounds and degree of any deformity. During the ‘feel’ stage the individual bones of the ankle are palpated using compression techniques to assess for possible fractures. In addition the medial and lateral collateral ligaments should be palpated to assess for tenderness, as any tenderness to the medial collateral ligament indicates a significant soft tissue injury. As the final part of the ‘feel’ assessment the clinician should assess the neurovascular status of the limb to check for potential neurovascular compromise distal to the injury.

The clinician then needs to assess the patient’s range of movement by asking the patient to move their foot through the planes of inversion, eversion, plantar and dorsi flexion, to assess ligament stability (Bickley 2003).

After completing the clinical examination the clinician should have an initial impression to the extent of the injury by taking into account the mechanism of injury and the examination findings. From these findings the clinician will decide whether the patient can be diagnosed with a sprain and discharged home without further investigation or whether an x-ray needs to be obtained to confirm the presence of a fracture.
2.6: Ankle X-rays

If an x-ray of the ankle is requested it involves the patient having an anteroposterior (AP) and a lateral x-ray. This enables the ankle to be viewed from the front and the side. Having the affected limb viewed in two images perpendicular to each other allows evaluation of displacement, angulation and rotation of the fracture site (Long & Rafert 1995). X-ray exposure should also enable evaluation of associated soft tissue injury and should allow evaluation of skin integrity and displacement of fat pads (Long & Rafert 1995). Figure 10 shows a typical AP and lateral ankle x-ray.

![Figure 10: Plain AP & lateral ankle x-rays](http://www.clinica-sandalf.com/Pictures/X-ray%20normal%20ankle.jpg)

According to Wall (1999) ankle x-rays account for 10% of all x-rays taken within an ED and Bachmann et al (2003) identified that only 10-15% of those receiving an ankle x-ray are diagnosed with a fracture. At a local level similar findings,
from an audit in an ED on the south coast of England in 2006, identified that only six out of 42 patients sent for an ankle x-ray had a fracture confirmed on x-ray. Sending a patient for an x-ray unnecessarily results in increased costs for the NHS in resources for taking the x-ray, and although the radiation from an ankle x-ray is deemed to be negligible some patients are ‘repeat’ attendees to ED and unnecessary x-rays add to an increase in exposure of radiation for these patients. Anis, Stiell, Stewart and Laupacis (1995) and Clements (2004, cited in Dissmann & Han, 2006) found that sending a patient to x-ray increased their length of stay by 36 and 31 minutes respectively. This will therefore have an impact on the 4-hour national standard set by the Department of Health [DH] for length of stay in an ED (DH 2001). In addition, sending a patient to the x-ray department unnecessarily impacts on the flow to the x-ray department for other patients. In order to address the problem of requesting unnecessary ankle x-rays Stiell et al (1992) devised the Ottawa Ankle Rules (OARs) in the early 1990s.

2.7: Development of the Ottawa ankle rules

In developing the OARs 750 adult patients were systematically assessed against 32 clinical variables associated with fractures to the foot and ankle (Stiell et al 1992). From this initial study Stiell and colleagues developed the first ankle and foot clinical decision rules. The aim of the rules was to identify adult patients who required ankle or foot x-rays to diagnose a clinically significant fracture. Stiell et al (1992) defined a clinically significant fracture to be anything other than small avulsion fractures, less than 3mm in depth, stating
that these fractures rarely need any form of management, other than self-care advice, and can be managed as simple sprains.

At the end of this first study Stiell et al (1992) concluded that an ankle x-ray was required if the patient

- Complained of pain near the malleoli
- Was 55 years or over
- Was unable to bear weight both immediately and in the ED
- Had bony tenderness to the posterior edge or tip of the malleoli

In 1993 Stiell and colleagues refined the rules and published the results of a second larger study which comprised two stages of development between February 1991 and January 1992 (Stiell et al 1993). 1032 adult patients presenting to EDs within 10 days of an ankle injury were assessed against 15 clinical variables associated with fractures to the ankle and foot. Stiell et al (1993) define the ankle to be ‘the area involved in twisting injuries’ (p. 1123) and subdivided it into malleolar and foot zones. The malleolar zone was defined as the area including the distal 6cm of the tibia and fibula and the talus. Sensitivity is reported as 100% (95% CI 97-100). Figure 11 shows a diagrammatic representation of the revised ankle and foot rules developed by Stiell et al (1993).
After refinement of the rules the second stage of the Stiell et al (1993) study assessed 453 patients against only 6 variables associated with fractures to the ankle and foot. 50 (11%) patients had fractures to the malleoli. The sensitivity remained high at 100% (95% CI 93-100) with a slight increase in specificity from 39% to 41% on part one of the study. Stiell et al (1993) concluded that the revised version of the ankle rules were accurate at diagnosing clinically significant fractures to the ankle in adult patients. The age criterion (previously set at 55 years) was removed altogether, and the weight bearing criterion was stated as being the most reliable criterion. Stiell et al (1993) conclude that an x-ray of the ankle is indicated in the presence of pain near the malleoli and

- The inability to weight bear both immediately or in the ED or
Bone tenderness at the posterior edge or tip of either malleoli or the distal fibula shaft (Stiell et al 1993).

Since their original development some twenty years ago the OARs have become one of the most studied and validated clinical decision tools used in clinical practice, and although they were initially designed for use in adults they have also been validated in paediatric populations. A Cochrane meta-analysis by Bachmann et al (2003) included studies on the use of the OARs in children and adults and identified 1085 citations from their search of Medline, Embase, the Cumulative Index to Nursing and Allied Health Literature [CINAHL] and the Cochrane library. In selecting the papers for inclusion Bachmann et al (2003) stated that they had to include sufficient data to produce a 2x2 table. Bachmann et al (2003) excluded papers that did not collect data prospectively and those where it was unclear whether the radiologist interpreting the x-ray had been blinded to the results of the OARs.

Twenty-seven studies met their inclusion criteria, 12 on the ankle, 8 on the foot, and 10 on the global accuracy of the rules. Bachmann et al (2003) included 15,581 patients, 47 (0.3%) were reported to be OARs negative but x-ray positive. The pooled sensitivity of all the studies was 98% (95% CI 96-99) with a median specificity of 32% (interquartile range 24-44). The 12 studies that focused on ankle assessment alone included 5945 patients and the pooled sensitivity for these was also 98% (95% CI 96-99) but a slightly higher median specificity at 40% (interquartile range 28-48). Bachmann et al (2003) did not
state the total number of fractures diagnosed in their analysis but concluded that less than 2% of patients who were OARs negative had a fracture on x-ray. It is not clear why Bachmann et al (2003) quote CI for sensitivity and Interquartile range for specificity.

In addition, a Cochrane systematic review by Dowling et al (2009) focused on the use of the OARs in patients aged less than 18 years of age. Dowling et al (2009) performed a search of Medline, Ovid Medline, Cochrane, Embase and CINAHL and identified 451 citations. 102 were selected for review with only 12 meeting the inclusion criteria. These 12 papers involved a total of 3130 patients. The mean age is not given but the range is quoted as two to seventeen years. Only 84 were under the age of five years. 671 fractures were identified, equating to a fracture prevalence of 21%. In similarity to the meta-analysis by Bachmann et al (2003) the pooled analysis in the Dowling et al (2009) paper revealed a sensitivity of 99% (95% CI 97-99).

In the original work Stiell et al (1993) reported the ability to bear weight criterion the most reliable in adults but Dowling et al (2006) raised concerns about its reliability in children. Dowling et al (2006) claimed a child’s reluctance to walk on a painful foot invalidated the weight bearing criterion in children and the specificities in the papers reviewed varied from 8% to 50%. Despite performing subgroup analysis Dowling et al (2009) stated they were unable to pool the specificities. The OARs missed ten fractures in the papers reviewed by Dowling et al (2006), equating to a false negative rate of 1.2% (95% CI 1-2) and reflects
a similar false negative rate to Bachmann et al (2003). Five of the missed fractures were from one study by Clark & Tanner (2003). Clark and Tanner (2003) included all fractures as being clinically significant in their study, claiming that regardless of the fracture type all have the potential to affect bone growth in children and should therefore all be deemed clinically significant. This is in contrast to the original work on the OARs which deemed small avulsion fractures less than 3mm in depth as not clinically significant in adults (Stiell et al 1993). However, the OARs were designed for use on adults and not children and as such did not include fractures associated with children in their definition of a clinically significant fracture. Although complications from both Salter-Harris I and Salter-Harris II fractures are rare they both need immobilisation for a short time to allow the injury to settle and Clark and Tanner (2003) therefore argue they should also be treated as clinically significant.

2.8: Implementation of the Ottawa ankle rules

The systematic review by Bachmann et al (2003) concluded that the implementation of the OARs is dependent on the level of sensitivity and specificity clinicians are prepared to accept, and ten years after their implementation Brehaut, Stiell, Visentin and Graham (2005) undertook a postal questionnaire to examine the use of the OARs by Canadian emergency physicians. The majority of the 376 eligible respondents rated the OARs favourably, with 252/376 (96%) rating them easy to use and 350/376 (93%) rating them useful in practice. Although 327/376 (87%) of the physicians stated
they remembered the rules Brehaut et al (2005) report that only 117/376 (31%) of the physicians could recite them in full.

Gravel, Roy and Carrière (2010) agreed with Brehaut et al (2005) stating that retention of the OARs was questionable and developed a mnemonic to try and aid retention of the rules amongst 206 doctors on a paediatric emergency medicine rotation. In this study the doctors were randomised to receive either an information sheet with the description of the OARs (control) or an information sheet containing the mnemonic '44-55-66-PM'. Figure 12 summarises the mnemonic.

<table>
<thead>
<tr>
<th>Mnemonic 44-55-66-PM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient needs an x-ray only if</td>
</tr>
<tr>
<td>4 - Unable to do 4 steps immediately AND</td>
</tr>
<tr>
<td>4 - Unable to do 4 steps in the ED</td>
</tr>
<tr>
<td>or</td>
</tr>
<tr>
<td>5 - has pain at the base of the 5th metatarsal</td>
</tr>
<tr>
<td>5 - Has pain at the Scaphoid of the foot (the navicular)</td>
</tr>
<tr>
<td>or</td>
</tr>
<tr>
<td>6 - Tenderness in 6cm Posterior edge of lateral Malleolus</td>
</tr>
<tr>
<td>6 - Tenderness in 6cm Posterior edge of medial Malleolus</td>
</tr>
</tbody>
</table>

Figure 12: Mnemonic 44-55-66-PM to aid retention of the OARs

Both groups were reported to have had similar knowledge of the OARs at the start of the study and also after three weeks. However, the long term retention
of the OARs was greater in the mnemonic group than in the control group and Gravel et al (2010) conclude that the mnemonic helped with retention the OARs. Specific to the UK Graham et al (2001) identified that 268/295 (91%) of doctors who were members of the British Association of Emergency Medicine [BAEM] were aware of the OARs, and 239/295 (81%) used them most of the time. In addition to the UK and Canada the OARs have also been successfully validated in France (Auleley et al 1998), Greece (Papacostas, Malliaropoulos, Papadopoulos & Liouliakis 2001), Australia (Bromfield & Stuart 2003), New Zealand (Wynn-Thomas et al 2002), Iran (Yazdani, Jahandideh & Ghofrani 2006), and the Netherlands (Knudsen, Vijdea & Damborg 2010).

However, Bachmann et al (2003) identified that the usefulness of the OARs had not been fully assessed within primary care and that the accuracy of the rules may be affected by the patient’s perception and acceptance of pain, which can be influenced by cultural background and previous experience to pain. In addition, Rosin and Sinopoli (1999) have urged caution in the implementation of the OARs within military populations due to the extent of physical exercise undertaken.

Bessen, Clark, Shakib and Hughes (2009) audited the use of the OARs in an Australian teaching and community hospital and revealed that they were either not being used or their application varied considerably. Individual, social and organisational barriers to the implementation of the OARs were identified, including lack of knowledge amongst clinicians about the OARs, lack of
confidence in the clinician’s ability to exclude a fracture without an x-ray, and the perception amongst the clinicians that ankles are not important. Bessen et al (2009) also reported that clinicians held the perception that patient’s expect an x-ray when they attend the ED, and that it is better to carry out an x-ray otherwise the patient will re-present to another healthcare provider. In support of this Wilson, Noseworthy, Rowe and Holroyd (2002) states that 78% of physicians allow the expectations of the patient to influence their compliance with the OARs and Kerr et al (1994) adds that as the patient has the right to be a partner in the decision making process they also have the right to request an x-ray.

2.9: Operator experience

Bachmann et al (2003) concluded that the accuracy of the OARs was affected by the experience of the operator using them. When the OARs were developed they were devised for use by experienced emergency physicians (Stiell et al 1993) but Mann, Grant, Guly and Hughes (1998), and Derksen et al (2005) have shown that the rules are also accurate when implemented by experienced ED nurses, reporting sensitivities of 98%, and 93% and specificities of 32% and 49% respectively.

However, in keeping with Bachmann et al (2003), Kerr et al (1994) and Perry, Raby and Grant (1999) have recommended caution when the OARs are used by inexperienced clinicians. Kerr et al (1994) reviewed the use of the OARs by physicians with varied emergency care experience and Perry et al (1999)
included clinicians who had not received any training on their use. Confidence intervals were not reported in either study but the sensitivities are quoted as being 94% for the Perry et al (1999) study and 93% for Kerr et al (1994). Kerr et al (1994) concluded that the financial savings of implementing the rules would be minimal and that the false negative rates were unacceptable within their study populations. It is not clear in these studies as to whether they used the same classification of a clinically significant fracture as Stiell et al (1993).

2.10: Reduction in x-rays

Stiell et al (1994) claim that within two years of their development the OARs reduced the number of ankle x-rays by 35%. Despite supporting the use of the OARs Mann et al (1998), Wynn-Thomas et al (2002) and Gwilym, Aslam, Ribbans and Holloway (2003) report a smaller reduction in the number of ankle x-rays after implementation of the OARs by 20%, 16% and 15% respectively.

In contrast studies by Cameron and Naylor (1999) and Holdroyd et al (2004) have reported an increase in ankle x-rays after implementation of the OARs. Following their 2-year prospective cohort study Holdroyd et al (2004) reported an increase of 1% in ankle x-rays after dissemination of the OARs amongst emergency physicians. In addition, Cameron and Naylor (1999) undertook a quasi-experimental before-and-after analysis following dissemination of the rules in 20 hospitals in Canada. Cameron and Naylor (1999) compared three groups of hospitals.
Chapter 2: Background & rationale for the study

- Firstly, Group A, included ten hospitals where the OARs were not routinely used prior to the study but who received dissemination of the rules;
- Secondly, Group B were five hospitals where the rules were used inconsistency before dissemination took place;
- Finally, Group C was used as the control and consisted of five hospitals where the rules were already used widely and staff received no additional dissemination of the rules.

Cameron and Naylor (1999) reported an increase in x-rays of 2% from group A and 8% for group B. This is compared to a 10% reduction for group C over the same time period. Cameron and Taylor (1999) concluded that local implementation in association with national dissemination techniques are required to encourage clinicians to use new guidelines.

2.11: Effect of the Ottawa ankle rules on waiting times

The OARs were developed for use by experienced clinicians at the time of clinical assessment but with the focus on reducing ED waiting times (DH 2001). Allerston and Justham (1997) and Fan and Woolfrey (2006) set out to establish whether a patient’s length of stay in the ED could be reduced if the OARs were performed by ENPs at the time of triage. Both these studies reported a reduced length of stay, but only Allerston and Justham (1997) found that the reduction was clinically significant from 81 minutes for patients sent to x-ray direct from triage compared to 107 minutes for the control group (p=0.001). This is
compared to the Fan and Woolfrey (2006) study where the difference between the two groups was only 7 minutes.

A unique attempt at using the OARs to reduce hospitals waiting times was tried by Blackman, Claridge and Benger (2008) who set out to identify whether patients attending an ED in the UK could be trained in the use of the OARs and apply them to themselves. They argued this would therefore reduce demand on busy healthcare systems. Blackman et al (2008) instructed patients in the use of the OARs and asked them to rate whether they felt they were OARs negative or positive against the different criterion. The patient self-reports were then compared to the clinician assessment. 100% of patients (50/50) rated themselves as OARs positive compared to only 90% (45/50) of clinicians (Blackman et al). Seven fractures were confirmed on x-ray and none were missed by the clinicians. Patients and clinicians had good agreement on their ability to assess bone tenderness with the most disagreement occurring with the weight-bearing criterion, where the clinician and patients only had agreement 17 times (Blackman et al 2008). This study was small scale and had no sample size calculations but Blackman et al (2008) concluded that educating patients on the use of the OARs had the potential to increase the demand on healthcare services rather than reduce it.

2.12: Alternative ankle rules

The meta-analysis by Bachmann et al (2003) and the systematic review by Dowling et al (2009) showed that the OARs have a consistently high sensitivity
but at the expense of a somewhat varied low specificity. Attempts to improve the specificity of the OARs have resulted in the development of other ankle assessment tools, and these include the

- Buffalo rules (Leddy, Smolinski, Lawrence, Snyder and Priore 1988)
- Leiden rules (Glas et al 2002; Pijenburg et al 2002)
- Utrecht ankle rules (Pijenburg et al 2002)
- Malleolar Zone rules (Dayan et al 2004) and
- Low-Risk Exam rules (Boutis et al 2001; Gravel, Hedrei, Grimard & Govin 2009).

Each of these is briefly described below.

2.12.1: Buffalo rules

Leddy et al (1998) reported the use of a modified version of the OARs. In an attempt to improve the specificity of the original OARs Leddy et al (1998) developed The ‘Buffalo’ rules, with the aim of using them in community sports medicine facilities. Leddy et al (1998) suggested that the OARs are not suited to the different fracture types that present to non-ED settings, claiming that the fractures are more subtle than those which present to EDs. The Buffalo rules involve palpating the crest of the malleoli, rather than the posterior edge so that the clinician is palpating away from the insertion points of the ligaments. Leddy et al (1998) reported a specificity of 42% (95% CI 31-55) for the OARs and 59% (95% CI 47-71) for the Buffalo rules, stating that x-rays could be reduced by
25% when compared to the OARs. The results are based on 132 patients, and despite these encouraging results no other papers were found that analyse the use of these rules.

### 2.12.2: Leiden & Utrecht rules

Both Glas et al (2002) and Pijnenburg et al (2002) compared the OARs with the Leiden ankle rules whilst Glas et al (2002) compared them to ‘physician clinical judgement’ and Pijnenburg et al (2002) to the Utrecht ankle rules. The Leiden and Utrecht ankle rules allocate a numerical value to the predictor variables of a fractured ankle, namely deformity, instability, crepitus, inability to bear weight, pulseless, pain on palpation, swelling and patient age. Patients are scored against each criterion with the total score indicating whether an x-ray is required or not. A score of seven for the Leiden rules, and a score of eight for the Utrecht rules indicates that an x-ray is needed. Table 1 summarises the criteria for the Leiden and Utrecht rules.

The Leiden rules were developed in the city of Leiden by Kievit, Dijkgraaf, Zwetsloot-Schonk, Tholen and Rapport (1991 – cited in Glas et al 2002). Glas et al (2002) included 647 adult patients presenting to an ED with ankle injuries and reported that the OARs were the most sensitive rule at 89% (95% CI 80-95), and physician clinical judgement the most specific at 80% (95% CI 74-87).
Table 1: Predictor variables of a fracture according to the Leiden & Utrecht ankle rules

<table>
<thead>
<tr>
<th>Leiden Rules</th>
<th>Score</th>
<th>Utrecht rules</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deformity, instability or crepitation</td>
<td>5</td>
<td>Deformity, instability or crepitation</td>
<td>4</td>
</tr>
<tr>
<td>Inability to bear weight</td>
<td>3</td>
<td>Inability to bear weight or axial compression</td>
<td>2</td>
</tr>
<tr>
<td>Pulseless or weakened posterior tibial artery</td>
<td>2</td>
<td>Pulseless or cyanosis</td>
<td>3</td>
</tr>
<tr>
<td>Pain on palpation of malleoli or 5&lt;sup&gt;th&lt;/sup&gt; metatarsal</td>
<td>2</td>
<td>Pain on palpation &amp; swelling to:</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Tibia</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Fibula</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Base 5&lt;sup&gt;th&lt;/sup&gt; metatarsal</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Achilles tendon</td>
<td></td>
</tr>
<tr>
<td>Swelling of malleoli or 5&lt;sup&gt;th&lt;/sup&gt; metatarsal</td>
<td>2</td>
<td>Haematoma or arthrosis</td>
<td>1</td>
</tr>
<tr>
<td>Swelling or pain of Achilles tendon</td>
<td>1</td>
<td>Age divided by 10</td>
<td>1</td>
</tr>
<tr>
<td>Age divided by 10</td>
<td></td>
<td>Total &gt; 8 x-ray requested</td>
<td></td>
</tr>
</tbody>
</table>

The OARs, Leiden rules and physician clinical judgement all missed fractures (8, 15, and 13 respectively). When adjusted to only include clinically significant fractures the OARs and physician clinical judgement missed one fracture each, compared to the Leiden rules, which missed five fractures (Glas et al 2002)

The clinicians involved in this study were junior surgical and orthopaedic registrars and Glas et al (2002) concluded that they did not need clinical decision rules to determine the need for ankle x-ray stating that clinical judgement alone would have considered x-rays for only 38% of patients, compared to 76% for the OARs and 46% for the Leiden rules. However, in this
Chapter 2: Background & rationale for the study

study all patients went on to receive an x-ray and Glas et al (2002) correctly identified that a limitation of the study is that the physicians involved knew in advance that all the patient in the study were to receive an x-ray regardless of their final clinical judgement and this pre-knowledge is likely to have biased the results.

The Leiden rules therefore appear to be less accurate than the OARs and this has also been supported in the study by Pijnenburg et al (2002) who scored all patients against the OARs, the Leiden rules and the Utrecht ankle rules. The sensitivity of the OARs was 98% (95% CI 87-100), compared to 88% (95% CI 74-98) for the Lieden rules and 59% (95% CI 42-74) for the Utrecht rules. The specificity was 26% (95% CI 22-29), 57% (95% CI 53-61) and 84% (95% CI 81-87) respectively. Pijnenburg et al (2002) concluded that although the Utrecht and Leiden rules resulted in the greatest reduction in the need for an x-ray this was at the expense of missing 41% and 12% of fractures respectively, stating that this is clearly unacceptable for clinicians and patients despite the potential reduction in x-ray costs. Table 2 summaries the results of the Pijenburg et al (2002) and the Glas et al (2002) studies.


<table>
<thead>
<tr>
<th>Study (date)</th>
<th>OARs</th>
<th>Physician judgement</th>
<th>Leiden ankle rules</th>
<th>Utrecht ankle rules</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sens % (CI)</td>
<td>Spec % (CI)</td>
<td>Sens % (CI)</td>
<td>Spec % (CI)</td>
</tr>
<tr>
<td>Glas et al (2002)</td>
<td>89 (80-95)</td>
<td>26 (23-30)</td>
<td>82 (72-90)</td>
<td>68 (64-71)</td>
</tr>
<tr>
<td>Pijenburg et al (2002)</td>
<td>98 (87-100)</td>
<td>26 (22-29)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
2.12.3: Malleolar-Zone & the Low-Risk Exam rules for use in children

As previously stated the OARs were devised for use on adults but have also been used on children with varying results so Dayan et al (2004) adapted them for use in children and developed the ‘Malleolar Zone ankle rules’. In keeping with Dowling et al (2009) Dayan et al (2004) claim that a child is often reluctant to walk after an injury therefore invalidating the weight-bearing criterion of the OARs. After reviewing the examinations of 717 patients Dayan et al (2004) found that 91 had a clinically significant fracture and identified that the high-risk criterion for a fracture in children were -

- Tenderness at either malleoli and inability to walk 4 steps in the ED
- Ability to walk in the ED but with tenderness and swelling at either malleoli or
- Tenderness just proximal to the lateral malleolus only

Dayan et al (2004) reported that the Malleolar Zone rules had a sensitivity of 100% (95% CI 88-99) and specificity of 19% (95% CI 17-23) when used on their cohort of 717 patients who had a median age of 13 years (IQ range 10-16 years). Dayan et al (2004) then used the assessment reports from their cohort of patients to assess them hypothetically against the OARs and found that the sensitivity would have been lower at 91% with a specificity of 10%.

Gravel et al (2009) compared the OARs to the Malleolar Zone rules and also the Low-Risk Exam rules on 272 patients aged less than 16 years. The Low-
Risk Exam rules were first described by Boutis et al (2001) and are defined as pain or tenderness, with or without oedema or bruising below the joint line of the distal fibula or over the anterior and posterior talofibular and calcaneo-fibular ligaments. All other findings are described as high risk for clinically significant fracture. In their original study Boutis et al (2001) claimed that the Low-Risk Exam rules could have reduced x-rays by 63% compared to 12% for the OARs without missing any clinically significant fractures. However, in contrast to the work by Stiell et al (1993) Boutis et al (2001) included non-displaced Salter-Harris, buckle, and epiphyseal avulsion fractures as not being clinically significant.

In contrast Gravel et al (2009) concluded that neither the Malleolar Zone nor the Low Risk Exam rules was superior at diagnosing clinically significant fractures than the OARs. Gravel et al (2009) found that the Low Risk Exam rules had a sensitivity of 76% (95% CI 66-84), the Malleolar Zone rules 93% (95% CI 85-96) and the OARs 99% (95% CI 93-100). The specificities were reported as 56% (95% CI 49-63), 27% (95% CI 21-33) and 30% (95% CI 24-37) respectively. In contrast to Dayan et al (2004) Gravel et al (2009) concluded that the Malleolar Zone rules increased the requesting of x-rays by 4% and missed three significant fractures, and although the Low Risk Exam rules had a superior specificity than the OARs six (14%) clinically significant fractures were missed. As with the Pijnenburg et al (2002) study this rate of missed fractures would clearly be unacceptable for both clinicians and patients.
In conclusion, the majority of evidence suggests that the OARs are widely accepted as an aid to assess ankle injuries. Although the figure varies the OARs appear to reduce the number of x-rays taken in most studies with reductions of between 15 and 35% reported. In addition they have been shown to reduce a patient’s length of stay in ED by up to 26 minutes when implemented at triage (Allerston & Justham 1997). Although the rules are widely validated in adults and children (Bachmann et al 2003; Dowling et al 2009) and are quick and easy to learn retention by some has been problematic (Brehaut et al 2005). Other barriers to their implementation include a lack of confidence by clinicians to exclude a fracture without x-ray (Bessen et al 2009), the clinician’s perception that patients prefer an x-ray (Wilson et al 2002), a lack of local dissemination (Cameron & Taylor 1999) and the perception that ankle injuries are not important (Bessen et al 2009). Finally, it would appear that teaching patients themselves to apply the OARs would increase demand on healthcare systems rather than reduce it (Blackman et al 2008).

When compared to other ankle rules, with the exception of the Buffalo rules (Leddy et al 1998), the OARs are more accurate, however making a true comparison across the studies is difficult as they all use different classifications of what constitutes a clinically significant fracture in a child. Stiell et al (1993) defined a small avulsion fracture less than 3mm in depth as not being clinically significant but this was in relation to adults. However, Gravel et al (2009) also classified Salter-Harris I fractures, Dayan et al (2004) small avulsion fractures, and Boutis et al (2001) non-displaced Salter-Harris I and II fractures, metaphyseal buckle fractures and epiphyseal avulsion fractures as being not
clinically significant in children. Boutis et al (2001) and Dayan et al (2004) base their definition on what constitutes a clinically significant fracture in children on whether immobilisation is required, but Clark and Tanner (2003) argue that all fractures have the potential to cause growth deformity in children and state that for this reason alone all fractures in children should be treated as significant.

2.13: Ottawa ankle rules & the tuning fork

In addition to the other ankle rules another attempt at improving the diagnostic accuracy of the OARs has been attempted in the UK by Dissmann and Han (2006) using the OARs in conjunction with a tuning fork. Dissmann and Han (2006) describe the tuning fork test as placing the flat end of a vibrating C 128Hz tuning fork (figure 13) on the tip of the lateral malleolus and then on the distal fibula shaft 5-10 cm proximal to the point of maximal tenderness.

![Flat end tuning fork](http://www.baileyinstruments.co.uk/UserFiles/Images/CH128.jpg)
49 patients were included in the study, of these seventeen recorded a positive tuning fork test when the tuning fork was placed over the lateral malleolus. There were seventeen positive tuning fork tests when the tuning fork was placed on the tip of the lateral malleolus and seven when the tuning fork was placed over the distal fibula shaft. Five patients had a fracture confirmed on x-ray, all of which had been identified by a positive tuning fork test at both sites. Dissmann and Han (2006) report a sensitivity of 100% (95% CI 46-100) and a specificity of 61% (95% CI 46-71) when the tuning fork was applied to the lateral malleolus and a sensitivity of 100% (95% CI 46-100) and specificity of 95% (95% CI 83-99) when applied to the distal fibula shaft. Importantly there were no false negatives and Dissmann and Han (2006) conclude that the potential reduction in x-rays was 66% and 86% respectively.

The Dissmann and Han (2006) study is well conducted and the results look promising however there are a number of limitations that need to be investigated further. Dissmann and Han (2006) have maintained a high sensitivity and increased the specificity compared to when the OARs are used in isolation but they report wide confidence intervals for sensitivity (46 – 100%) and therefore the study does not have the necessary sample size to detect the effect. Although the results of the tuning fork test and x-ray were reported blind of each other the patients were only recruited if the bony prominence of the lateral malleolus was palpable (i.e. not hidden by swelling) and the tuning fork test was performed by a single operator limiting practicalities of implementing the tuning fork test into busy emergency care environments.
If ankle injuries account for 8% of all minor injuries and 10% of all x-rays taken in an ED (Wall 1999) even the lower 66% reduction in x-rays in the Dissmann and Han (2006) study would amount to a significant saving to the NHS. The total cost to the NHS for ankle x-rays in the UK is unknown, but Fiesseler, Szucs, Kec and Richman (2002) report that in Canada and the USA the cost was an estimated $500million annually when they undertook their study ten years ago. Anis et al (1995) undertook a before and after intervention study and compared the proportion of patients referred for x-ray before and after the implementation of the OARs and found that x-rays were reduced by 28% (95% CI 22-33). Anis et al (1995) then performed a cost effective analysis on the same two cohorts of patients and reported that when all the costs involved in obtaining an x-ray, that is the cost of the x-ray, the physician costs, the average hourly wait, the cost of time saved in ED, loss of wages, and litigation settlement there could be a cost saving per 100,000 patients in the US of between $614,226 to $3,145,910, and in Canada of $730,145 after implementation of the OARs. This study was reported some 16 years ago so it is reasonable to assume that healthcare costs and patient expectations have risen dramatically and that if repeating the study today these costs would be much higher. In light of the extensive cost cutting exercises now being seen in the health service in the UK any reduction in costs to the NHS is paramount and needs to be explored.

Could a simple adjunct like the tuning fork test be acceptable to patients and clinicians, and could it increase the specificity of the OARs whilst maintaining the high sensitivity when used to assess twisting ankle injuries in multiple
emergency care settings by multiple operators? In order to answer this question a literature review was undertaken to identify what is already known about the use of the tuning fork in detecting fractures and this can be found in the next chapter.
Chapter 3 – A SYSTEMATIC REVIEW OF THE LITERATURE ON THE USE OF THE TUNING FORK TO DETECT FRACTURES

This chapter will summarise the use of the tuning fork and identify what is already known about its application in the detection of fractures and outline the research question and hypothesis.

3.1: History of tuning fork

In 1550 Italian physician, mathematician and astrologer Cardano described how sound waves pass through the skull but it was a few years before Capuacci realised the phenomena might be useful as a diagnostic test to differentiate between different types of hearing loss (Feldman 1997a). In 1684 German physician Schelhammer used a common dinner fork to further explore this phenomenon (Feldman 1997a).

Later in 1711 John Shore, who was trumpeter and lutenist to Purcell and Handel in London, invented the tuning fork for use as a musical instrument and in 1825 Ernest Webber discovered that the vibration emitted from a tuning fork could be used to differentiate between conductive and sensorineural hearing loss (Feldman 1997a). The ‘Webber’ test as it later became known, consisted of placing a vibrating tuning fork in the middle of the head (Davitt, n.d.a). The patient was then asked whether they could hear the sounds better in one ear or both the same. In a ‘normal’ Webber test the sound is symmetrical, if the sound localises to the ‘poor’ ear conductive hearing loss is diagnosed, and if the
sound localises to the ‘good’ ear sensorineural hearing loss is diagnosed. Despite this breakthrough the Webber test was not introduced into medical practice until 1845 by otologist Schmalz (Bickerton and Barr 1987).

Later in 1855 Heinrich Adolf Rinne again demonstrated the difference between conductive and sensorineural hearing loss by placing a vibrating tuning fork on the mastoid bone to detect bone conduction (Davitt, n.d.b). The patient is asked to tell the examiner when the sound is no longer heard and the tuning fork is then moved immediately in front of the ear, to test air conduction, and again the patient asked to inform the examiner when the sound diminishes. The time interval of each is noted and in a ‘normal’ Rinne test the air conduction lasts twice as long as bone conduction. In conductive hearing loss bone conduction is heard longer than air conduction, and in sensorineural hearing loss air conduction is heard longer than bone conduction in the affected ear. It was another 25 years before the ‘Rinne’ hearing test was introduced into routine medical practice (Feldman 1997b).

Both the Rinne and Webber hearing test identified that the sound waves emitted from a vibrating tuning fork transmit through bone but in order to identify evidence of its use in detecting fractures a review of the literature review was undertaken.
Chapter 3: A systematic review of the literature on the use of the tuning fork to detect fractures

3.2: Review method

The literature search initially took place in October 2006, and was updated in 2011.

3.2.1: Search strategy & results

The CINAHL, Medline, SPORTSDicuss, Medion, and the Allied and Complimentary Medicine (AMED) databases were searched. CINAHL is an authoritative resource for nursing and allied health professionals, Medline is an on-line database of journals focusing on life sciences, and the SPORTSDiscuss database focuses on journals from the fields of fitness and sport. In addition an extensive examination of current literature using the internet search engine Google was undertaken. The AMED database was included as I thought that the tuning fork may have been used in the field of complementary medicine as well as conventional medical practice. A description of the databases and the dates searched can be found in table 3. No language or date limits were set except for the availability of the databases on-line. Due to the small number of papers expected the search was not limited to diagnostic research papers but included other primary designs, for example, case studies, and secondary synthesis including literature reviews, narratives and editorial comments.
Table 3: Databases searched - description & dates

<table>
<thead>
<tr>
<th>Database &amp; date searched</th>
<th>Brief Description</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>CINAHL – Cumulative Index to Nursing and Allied health Literature</td>
<td>Authoritative resource for nursing &amp; allied health professionals, students, educators and researchers. Focuses on UK and worldwide nursing and allied health subjects from more than 3000 journals dating back to 1981</td>
<td>1998-2011</td>
</tr>
<tr>
<td>Medline</td>
<td>The Medical Literature Analysis and Retrieval System, 21 million records from 5000 publications focusing on the life sciences and biomedical information</td>
<td>1950-2011</td>
</tr>
<tr>
<td>SPORTSDiscuss</td>
<td>Provided by the Sports Information Resource Centre and offers comprehensive, bibliographic coverage of sports, fitness and related disciplines. Dates back to 1800s.</td>
<td>1992-2011</td>
</tr>
<tr>
<td>Medion (<a href="http://www.mediodatabase.nl/">http://www.mediodatabase.nl/</a>)</td>
<td>An on-line collection of systematic reviews focusing on diagnostic and screening studies</td>
<td>2003 - 2011</td>
</tr>
<tr>
<td>AMED – Allied and complementary medicine</td>
<td>Contains a bibliography of literature relevant to the disciplines of complementary medicine, as well as physiotherapy and occupational therapy,</td>
<td>1985 - 2011</td>
</tr>
<tr>
<td>Google</td>
<td>Widely used internet search engine that uses text matching techniques to find web pages No date was set but the search ended when the citations no longer appeared relevant to the topic</td>
<td>No date</td>
</tr>
</tbody>
</table>

The initial search in 2006 searched the abstracts of Medline (1950 to 2006) and CINAHL (1998 to 2006) with the free-text term ‘tuning fork’. This search
revealed that the tuning fork had also been used in a technique called ‘auscultatory percussion’ or ‘osteophony’ so a further search using the free text term ‘auscultatory percussion’ or ‘osteophony’ [AND] ‘fracture/s’ was undertaken using the same databases. This resulted in a total of 220 citations being identified. Three of these were duplicates, leaving 217 citations for review.

Assessment of the abstracts identified that the majority focused on assessment of hearing loss and the detection of neuropathy due to diabetes, chemicals and drugs. Only eight citations focused on the use of the tuning fork to detect fractures, and one of these was the Dissmann and Han (2006) study. The others included research papers from Misurya, Khare, Mallick, Sural and Viswakarma (1987), Adams, Yarnold and Matthews (1988), Lesho (1997), Adams and Yarnold (1997), and Tiru, Goh and Low (2002), a university thesis by Moore (2005), a literature review by Kazemi (1999), and case studies by File, Wood and Kreplick (1996), Minardo (1997), and Gleberson & Hyde (2006).

A second search of the CINAHL, Medline, Medion, and AMED databases in 2011 using the free-text terms ‘tuning fork’ [AND] ‘fracture/s’, ‘osteophony’ [OR] ‘auscultatory percussion’ [AND] ‘fractures’ was used to search the abstracts and titles of the citations. In this instance the date was set from 2006 to 2011 to prevent duplication of the citations identified in the original search. This resulted in 132 hits, of which 10 were duplicates identified in the original search. Only three new citations were identified that focused on the tuning fork in the
detection of fractures (Wilder, Vincent, Stewart, Pack & Vincent 2009; Moore 2009; and Rapp 2009). In addition an extensive search of ‘Google’, using the same free text terms, was undertaken resulting in a further four studies being identified (McGaw 1942; Waldron & Hurley 1988; Tuling 2000; Van den Berg 2003). Figure 14 summarises the search results.

Figure 14: Summary of search results

A total of 17 papers were identified as being relevant to the topic. Full text versions of these papers were obtained. On review the Moore (2009) paper was a peer review updated publication of the thesis identified in the first search (Moore 2005), so the original thesis was excluded. Review of the reference lists identified four further papers (Lippmann 1932; Colwill & Berg 1958; Peltier 1977; Bache & Cross 1984). A summary of the author, date, country of origin,
study setting, sample size and patient age of the primary research papers can be found in table 4.

<table>
<thead>
<tr>
<th>Author (date) &amp; Country of origin</th>
<th>Main aim</th>
<th>Study design</th>
<th>Study setting</th>
<th>Sample size</th>
<th>Patient age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bache &amp; Cross (1984) UK</td>
<td>To use a tuning fork and stethoscope to diagnose fractured neck of femur – the Barford test</td>
<td>Diagnostic test study</td>
<td>ED</td>
<td>100</td>
<td>Ave age 79y - range not reported</td>
</tr>
<tr>
<td>Misurya et al (1987) India</td>
<td>Comparison of clinical examination &amp; auscultation in fractures to the femur &amp; tibia</td>
<td>Diagnostic test study</td>
<td>Fracture clinic</td>
<td>50</td>
<td>not reported</td>
</tr>
<tr>
<td>Adams &amp; Yarnold (1988) UK</td>
<td>Evaluate olecranon-maebrium percussion in shoulder trauma</td>
<td>Diagnostic test study</td>
<td>ED</td>
<td>47</td>
<td>Not reported</td>
</tr>
<tr>
<td>Lesho (1997) USA</td>
<td>To evaluate sensitivity &amp; specificity of the tuning fork test relative to bone scanning in detection of tibial stress fractures</td>
<td>Diagnostic test study</td>
<td>Military clinic</td>
<td>46</td>
<td>Mean 25y (range 19-43)</td>
</tr>
<tr>
<td>Adams et al (1997) UK</td>
<td>Evaluate accuracy of patella-pubic percussion in hip fractures</td>
<td>Diagnostic test study</td>
<td>ED</td>
<td>41</td>
<td>Not known</td>
</tr>
<tr>
<td>Tiru et al (2002) Singapore</td>
<td>To study sensitivity &amp; specificity of auscultatory percussion in occult hip fractures</td>
<td>Single blind prospective study</td>
<td>ED</td>
<td>290</td>
<td>Mean 72y ±7</td>
</tr>
<tr>
<td>Van den Berg (2003) USA</td>
<td>To determine the ability of ultrasound (USS) and the tuning fork test to diagnose fractures</td>
<td>Diagnostic test study</td>
<td>Clinic</td>
<td>200</td>
<td>Mean 31y ±13</td>
</tr>
<tr>
<td>Dissmann &amp; Han (2006) UK</td>
<td>To determine the suitability of tuning fork testing for increasing the specificity of detecting fractures to the lateral malleolus when using the Ottawa ankle rules</td>
<td>Diagnostic test study</td>
<td>ED</td>
<td>49</td>
<td>Range 12-84y</td>
</tr>
<tr>
<td>Wilder et al (2009) USA</td>
<td>To examine the sensitivity &amp; specificity of 3 tuning forks to detect stress fractures</td>
<td>Diagnostic test study</td>
<td>Runners clinic</td>
<td>45</td>
<td>Mean 31y Range 18-31</td>
</tr>
<tr>
<td>Moore (2009) USA</td>
<td>To assess the diagnostic accuracy of a tuning fork &amp; stethoscope technique in detecting fractures in patients</td>
<td>Diagnostic test study</td>
<td>Athletic training room</td>
<td>37</td>
<td>Range 7-60y</td>
</tr>
</tbody>
</table>

A total of 21 papers were identified as being suitable for inclusion in the review, these contained a mix of primary research papers, literature reviews and narrative papers describing the technique of using the tuning fork or
auscultatory percussion to assess fractures. The main findings, sensitivity, specificity, strengths and weaknesses of the research papers identified can be found in Table 5.

Table 5: Main findings, sensitivity, specificity, strengths & weaknesses of diagnostic research papers

<table>
<thead>
<tr>
<th>Author (date)</th>
<th>Main findings</th>
<th>Sensitivity % (CI)</th>
<th>Specificity % (CI)</th>
<th>Strengths &amp; weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bache &amp; Cross (1984)</td>
<td>Detected 51 (91.1%) fractures but incorrect in 8 patients</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Consecutive patients. Sensitivity &amp; specificity not reported but able to calculate from results</td>
</tr>
<tr>
<td>Misurya et al (1987)</td>
<td>Tuning fork correct in 94% 3 false negatives</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Clear description of index test. No blinding. Unable to produce 2x2 table from results</td>
</tr>
<tr>
<td>Adams &amp; Yarnold (1988)</td>
<td>40 abnormal tests. Identified clavicle fractures (100%), humeral head fractures &amp; 90% dislocations.</td>
<td>Not reported</td>
<td>Not reported</td>
<td>p values reported as &lt;0.005. No 2x2 tables. Unsure if results were blinded</td>
</tr>
<tr>
<td>Lesho (1997)</td>
<td>Stress fractures diagnosed in 60% Conclude tuning fork is not accurate at ruling out stress fractures</td>
<td>75 (CI not reported)</td>
<td>67 (CI not reported)</td>
<td>Tuning fork test method described in detail. Data not easily displayed. Confusing data</td>
</tr>
<tr>
<td>Adams et al (1997)</td>
<td>Identified 15 out of 19 hip fractures in 41 patients</td>
<td>Not reported</td>
<td>Not reported</td>
<td>p value 0.0001</td>
</tr>
<tr>
<td>Van den Berg (2003)</td>
<td>83 false positives &amp; 15 false negative tuning fork test.</td>
<td>Tuning fork 22 USS 8</td>
<td>Tuning fork 78 USS 78</td>
<td>2x2 tables incorrectly displayed, confusing data.</td>
</tr>
<tr>
<td>Dissmann &amp; Han (2006)</td>
<td>Tuning fork test more sensitive when placed on distal fibula shaft (DFS) &amp; lateral malleolus (LM). No false negatives reported</td>
<td>100</td>
<td>LM = 61 CI reported as 46-75 but should be 46-76</td>
<td>Single operator, single site, small numbers of patients &amp; fractures included. Incorrect CI given for sensitivity and specificity</td>
</tr>
<tr>
<td>Wilder et al (2009)</td>
<td>256Hz Tuning Fork elicited highest pain rating</td>
<td>128Hz = 83 256Hz = 92 512Hz = 77</td>
<td>128Hz = 38 256Hz = 19 512Hz = 65</td>
<td>CI intervals given. Data easy to read. Not all patients received the same reference test</td>
</tr>
<tr>
<td>Moore (2009)</td>
<td>20 true negatives, 10 true positives, 5 false positives, 2 false negatives</td>
<td>83 CI not reported calculated as 57-97</td>
<td>80 CI not reported calculated as 59-93</td>
<td>No sample size calculations CI not reported</td>
</tr>
</tbody>
</table>
The case studies provide useful in-depth information on individual cases, but the small numbers involved do not provide reliable evidence that is replicable. The literature review and narratives again provide background information and useful secondary reference, but again do not provide reliable evidence of tuning fork diagnosis in the population of interest. The case studies and secondary synthesis articles on the use of the tuning fork and auscultatory percussion can be found in table 6.

**Table 6: Summary of case studies, literature review & narratives on auscultatory percussion & tuning forks in the detection of fractures**

<table>
<thead>
<tr>
<th>Author (date)</th>
<th>Main finding from paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lippmann (1932)</td>
<td>Narrative paper on the use of auscultatory percussion to identify position of fracture fragments to the hip, clavicle and humerus</td>
</tr>
<tr>
<td>Colwill et al (1958)</td>
<td>Narrative paper on the use of auscultation in the diagnosis of hip fracture in 50 patients</td>
</tr>
<tr>
<td>Peltier (1976)</td>
<td>Narrative paper on the use of auscultatory percussion to detect hip fractures</td>
</tr>
<tr>
<td>Waldron &amp; Hurley (1988)</td>
<td>Case study - Use of a tuning fork to diagnose temporal fractures in 8 male patients post head injury. Tuning fork test positive in all 8, fracture diagnosed in 5</td>
</tr>
<tr>
<td>File et al (1996)</td>
<td>Case study - Use of auscultatory percussion in 2 patients with normal plain x-rays. Auscultatory percussion positive and hip fracture diagnosed on scanning in both</td>
</tr>
<tr>
<td>Minardo (1997)</td>
<td>Case study - Positive tuning fork test to a hockey player with non-traumatic painful shoulder. X-ray confirmed presence of a bone tumour</td>
</tr>
<tr>
<td>Kazemi (1999)</td>
<td>An exploratory literature review of the use of tuning forks &amp; Ultrasound scan to detect fractures</td>
</tr>
<tr>
<td>Tuling (2000)</td>
<td>Case study - Tuning fork failed to detect radial head fracture in a female patient following a fall from a bicycle</td>
</tr>
<tr>
<td>Gleberzon &amp; Hyde (2006)</td>
<td>Case study - Tuning fork detected fracture to hip in patient with back pain after a fall down stairs</td>
</tr>
</tbody>
</table>
3.2.2: Discussion

Auscultatory percussion of bone or ‘osteophony’ is described as placing a stethoscope over a bony prominence and tapping the opposite end of the same bone with a finger and was first described as a technique for detecting fractures by Lisfranc in 1823 (Peltier (1977). Colwill and Berg (1958) states that along with crepitus osteophony is the only sign present after trauma that is directly dependent on a fracture rather than injury to the surrounding soft tissues. The use of osteophony in the detection of fractures is presented in the literature by Lippmann (1932), Colwill and Berg (1958), Peltier (1977), Adams et al (1988), Adams and Yarnold (1997), File et al (1998), and Tiru et al (2002).

According to Lippmann (1932) the sound heard during osteophony is loud, high-pitched and resonant and is a rapid painless indicator of the position of fracture segments and can be used to determine the extent of healing. This narrative account recommended that osteophony can be useful in detecting fractures to the femur, humerus and clavicle, and is particularly useful when x-ray facilities are either unavailable, or inadequate.

Colwill and Berg (1958) were the first to report the findings of osteophony on 50 patients with a confirmed fracture against a control group of patients without fracture. There were no false positives but six fractures were missed – one to the pubic ramus, one to an impacted neck of femur, one to a compression fracture of the tibial plateau and it was negative to all three patients with a fracture to the ankle. Colwill and Berg (1958) conclude that osteophony is
difficult to apply to the ankle due to the anatomy of the ankle joint and the
difficulty in percussing the fragments on either side of the fracture line.

Adams and Yarnold (1997) assessed the accuracy of osteophony in 41 patients
with suspected hip fracture post injury. All patients received an x-ray and 19
were confirmed with a fracture. 14 (79%) of these had recorded an abnormal
osteophony test (p. 0.0001). ‘Occult’ fractures are where there is a high index of
suspicison for a fracture but plain x-rays are normal and both File et al (1998)
and Tiru et al (2002) describe the use of osteophony to detect occult fractures
to the neck of the femur. File et al (1998) describe two cases where osteophony
was ‘positive’, despite normal x-rays, and in both cases fractures were later
confirmed with further imaging. Tiru et al (2002) studied 290 patients with
suspected fractured neck of femur who all had normal x-rays. Osteophony was
performed and was positive in 245 patients. All 290 patients went on to receive
either a bone scan, magnetic resonance imaging (MRI) or computerised
tomography (CT) and 245 patients were diagnosed with a hip fracture. Tiru et al
(2002) report sensitivity to be 96% (95% CI 87-99) and specificity 86% (95% CI
49-98) and conclude that as physical examination and x-rays are not 100% any
test that increases diagnostic accuracy should be considered to raise suspicion
for fracture. Tiru et al (2002) also recommend the use of osteophony when
assessing the unconscious, demented or uncooperative patient rather than
relying on clinical judgement alone. File et al (1998) add that in the presence of
a normal x-ray, but strong clinical suspicion of a fracture, further imaging should
be requested if osteophony is positive.
Adams et al (1988) used osteophony to detect abnormalities to the shoulder following injury and found that the test was accurate in all clavicle fractures (N = 9, p. <0.005), 16 of 20 humeral fractures (p. <0.01) and 11 of 13 shoulder dislocations (p. <0.02). They conclude however that osteophony is not accurate in the detection of acromio-clavicular joint abnormalities.

Bache and Cross (1984) were the first to describe the use of the tuning fork to diagnose fractures by modifying the osteophony technique. When using a tuning fork it is ‘thought that the fractured bone vibrates, resulting in irritation of the over-lying periosteum which evokes pain’ (Tuling 2000, p. 36). Osteophony using a tuning fork was first used by Dr. Barford of the East Birmingham Hospital in the diagnosis of fractures to the neck of femurs. Bache and Cross (1984) included 100 patients in their study, 48 were deemed to have a fracture when assessed using conventional clinical methods and 51 patients recorded a positive ‘Barford’ test. X-rays revealed that 56 patients actually had a fracture. The combination of clinical assessment and the Barford test missed three fractures, one patient was obese, another had a fractured pubic rami and another had an undisplaced fractured neck of femur. Bache and Cross (1984) conclude that when used with clinical examination the Barford test would assist GPs in diagnosing neck of femur fractures in the community therefore preventing unnecessary journeys to hospital for this vulnerable patient group.

Misurya et al (1987) and Moore (2009) used the ‘Barford test’ to assess for long bone fractures. As with the previous studies Misurya et al (1987) used a
paediatric stethoscope but Moore (2009) used the bell end of an adult stethoscope. In both studies the tuning fork was placed distal to the injury and the stethoscope proximally on the same bone. However, Moore (2009) modified his technique in the presence of swelling and placed the tuning fork and stethoscope in reverse order. Misurya et al (1987) included fifty patients and report three false negative results - two fractured necks of femur and one femoral shaft fracture and conclude that the test is accurate and would be useful in situations that result in mass casualties and in agreement with Tiru et al (2002) would be useful in the uncooperative and unconscious patient. Moore (2009) studied 37 patients attending an athletic training room. There were 10 true positive, 20 true negative, 5 false positives and 2 false negatives and Moore (2009) report that sensitivity was 83% and specificity 80%. Confidence intervals are not given but from the data reported in the paper these are calculated as 57-97 for sensitivity, and 59-93 for specificity. The tuning fork test was negative to the only avulsion fracture (5th metatarsal) and the only buckle fracture (clavicle) included in the study and the author concluded that although the test was simple and quick to administer as well as painless and inexpensive it may be unreliable in detecting buckle and avulsion fractures (Moore 2009).

Although papers focusing on the use of the tuning fork to assess hearing loss were excluded from this literature review Waldron and Hurley (1988) discuss the use of the Webber and Rinne tests to detect fractures to the temporal bone in eight patients post head injury. Three patients had been diagnosed with skull fractures to the occiput and parietal regions on x-ray, however there was high suspicion of a temporal bone fracture in all patients. The temporal bone
contains the middle and inner ear and these can become damaged in the presence of a fracture (Waldron & Hurley 1988) and the Rinne and Webber tests were abnormal in all eight patients and all went on to have further imaging. Five were diagnosed with a temporal bone fracture and Waldron and Hurley (1988) concluded that all patients presenting with head injuries should receive tuning fork testing to assess for potential temporal bone fractures.

Lesho (1997) was the first to use a tuning fork without a stethoscope to detect fractures after becoming concerned about the number of bone scans requested in a military sick bay for potential stress fractures. Stress fractures are caused by repetitive stress or strenuous exercise (Brukner, Bennell & Matheson 1999) and Lesho (1997) describes placing a 128Hz tuning fork over the anterior shin and moving it distally towards the ankle. A positive test was said to have occurred if the pain increased and was localised to an area less than 3cm in diameter (Lesho 1997). All patients (N=46) received the tuning fork test and a bone scan. The number of positive and negative tuning fork tests and bone scans is not reported but 60% were diagnosed with a fracture, this equates to 27 patients. Sensitivity was reported as 75% and specificity as 67% (Lesho 1997).

Only in the advanced stages of a stress fracture is there disruption to the periosteum of the bone and Lesho (1997) concluded that only in the later stages of a stress fracture would the fracture be positive to the tuning fork test. However, within this study there was a thirty day time lag between the tuning
fork test and the bone scan being performed and disease progression bias (Whiting, Rutjes, Reitsma, Bossuyt & Kleijnen 2003) could be introduced into this study. Lesho (1997) concluded by stating that the tuning fork test could prove useful to perform immediately before a bone scan as a positive result would negate the need for a scan, whereas a patient recording a negative tuning fork test should continue to have a scan. Wilder et al (2008) agree stating that when used on stress fractures further imaging is required when a negative tuning fork test is achieved.

In a descriptive review of the literature Kazemi (1999) focused on the use of the tuning fork and ultrasound scan in the detection of fractures and concluded that the validity and reliability of the tuning fork in the detection of acute fractures has not been established. Kazemi (1999) recommend that a validity study, using plain x-rays as the gold standard, is undertaken to assess the use of the tuning fork in the detection of fractures.

Since then Van den Berg (2003) have compared the tuning fork test with ultrasound in 200 patients with suspected fractures. Van den Berg (2003) report that they excluded patients with suspected scaphoid and femoral head fractures but do not include the sites of the body that were included. Furthermore, the data in the Van den Berg (2003) study is misleading and inconsistent. They include 23 patients who had ‘no bony tenderness’ and 14 who had ‘obvious deformity’ and I would argue that if there is no bony tenderness or obvious deformity on clinical examination assessment with a tuning fork is not
necessary. In addition, the 2x2 tables included within the results are incorrect as Van den Berg (2003) have the axis of their 2x2 tables labelled in reverse.

The narrative by Lippmann (1932) and the study by Moore (2009) both recommended that the tuning fork test could be useful in settings where x-rays are not readily available and Gleberzon and Hyde (2006) describe the case of a 54 year old patient presenting to a chiropractic clinic following a fall down stairs. The patient was complaining of back pain and when the tuning fork was applied to her hip she experienced severe pain. She was later confirmed to have a fracture to the hip on x-ray. Another case study described the use of the tuning fork on a patient complaining of a painful shoulder, there had been no history of trauma, but the tuning fork test was positive and x-ray later revealed a bone tumour (Minardo 1997). Tuling (2000) however, found that the tuning fork test was not sensitive enough to diagnose a radial head fracture in a patient following a fall from a bike. In this case study however Tuling (2000) describes placing the tuning fork on the bony prominences of the elbow, and I would argue that as the radial head is not a bony prominence good connection with the fractured bone may not have occurred in this case.

All the papers discussed so far use a 128Hz tuning fork as recommended for assessing vibratory sensation (Bickley 2003) but Wilder et al (2008) compared the results of three different sizes of tuning fork to assess for stress fractures in 45 athletes and then compared them to the results of either x-ray, MRI or bone scan. A summary of the results for each size tuning fork can be seen in table 7.
Table 7: Results of Wider et al (2008) comparing different tuning forks in the detection of stress fractures

<table>
<thead>
<tr>
<th></th>
<th>X-ray n=44</th>
<th>MRI n=16</th>
<th>Bone scan n=22</th>
</tr>
</thead>
<tbody>
<tr>
<td>128 Hz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>83.3</td>
<td>80.0</td>
<td>70.5</td>
</tr>
<tr>
<td>Specificity</td>
<td>37.5</td>
<td>50.0</td>
<td>60.0</td>
</tr>
<tr>
<td>256 Hz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>92.3</td>
<td>90.0</td>
<td>77.7</td>
</tr>
<tr>
<td>Specificity</td>
<td>19.3</td>
<td>20.0</td>
<td>25.0</td>
</tr>
<tr>
<td>512 Hz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>76.9</td>
<td>50.0</td>
<td>35.5</td>
</tr>
<tr>
<td>Specificity</td>
<td>64.5</td>
<td>83.3</td>
<td>40.0</td>
</tr>
</tbody>
</table>

Wilder et al (2008) identified that the 256Hz tuning fork produced the highest pain score during the test when compared to the 128Hz and the 512Hz but concluded that if a patient rated the pain as three on a three-point pain scale with any of the tuning forks it was highly suspicious of a fracture. No confidence intervals are given in the paper but the 256Hz is reported as being the most sensitive but least specific when compared to all three tests.

3.2.3: Conclusion

Whilst the use of the sound waves to detect fractures was first documented in the 1800s the evidence base for its diagnostic use remains uncertain. The literature reveals that there are mixed views as to whether the tuning fork test is accurate at diagnosing fractures and that it appears to be dependent on the site of injury, the type of fracture, as well as whether there is swelling, deformity, and even whether the patient is obese. It also appears that good contact with the bone is essential. Several of the papers recommend further research into the use of the tuning fork test to confirm or exclude its accuracy in the detection of fractures before it is introduced into practice and although I have personal
experience of its use in the military for the detection of stress fractures current
evidence suggests that a negative tuning fork does not negate the need for
further imaging, whereas a positive tuning fork test can.

Fractures to the ankle are one area where osteophony using a tuning fork test
has proven difficult due to the placement of the stethoscope. However, the
recent study by Dissmann and Han (2006) revealed that when used in
conjunction with the OARs the tuning fork test was accurate when used by a
single operator on ankles following inversion injury in the absence of swelling.
In order for the results of the Dissmann and Han (2006) study to be widely
disseminated the results need to be replicated when used by multiple
operators, with a variety of experience, in multiple emergency care settings.
The following section will outline a study that aims to assess the accuracy of the
tuning fork in detecting ankle fractures.

3.3: Research question & hypothesis

In light of the evidence identified in the literature two research questions will be
answered, the first will be a diagnostic question to assess the accuracy of the
tuning fork test when used in the detection of fractures by multiple operators,
and the second will be a psychosocial question in order to explore patient
experience and clinician acceptability of the tuning fork test. The two questions
will therefore be
1) “Can the tuning fork test increase the diagnostic accuracy of the Ottawa ankle rules when used by multiple operators on patients attending emergency care settings following twisting ankle injuries?”

2) “What is the experience of patients receiving and clinicians using the tuning fork and what do they view is the potential impact for patients if the test were introduced into clinical practice?”

The one-tailed hypothesis for question one above is that when the tuning fork test is used in conjunction with the OARs the diagnostic accuracy of the OARs will increase when used on patients attended emergency care settings following simple twisting ankle injuries.

The null hypothesis for question one above is that the tuning fork test will not change the diagnostic accuracy of the OARs when used on patients attended emergency care settings following simple twisting ankle injuries.

3.3.1: Outcome measures

The main outcome measure was to identify the diagnostic accuracy of the tuning fork test in conjunction with the OARs when used on simple twisting ankle injuries.

The secondary outcome measures include whether the accuracy of the tuning fork test is affected by the experience and role of the operator, and whether it is
affected by patient demographics, such as age, gender, ethnicity, degree of swelling, and distracting or previous injury. In addition, patient and clinician’s experiences will be sought in order to explore their acceptance on the use of the tuning fork test in clinical practice.

In order to answer the two research questions and to meet the study outcomes a two phase study involving both quantitative and qualitative methods has been undertaken.

3.3.1.1: Phase one – diagnostic test study

The primary objective of the study focuses on a diagnostic test study to assess the accuracy of the tuning fork test in the detection of ankle fractures. A diagnostic test is defined by Irwig, Bossuyt, Glasziou, Gatsonis and Lijmer (2002) as a test which seeks to confirm the presence of a disease in symptomatic patients. A diagnostic test allows patients to be classified into two groups, those with the target condition resulting in a positive test, and those without the target condition resulting in a negative test. Diagnostic tests have to be assessed for accuracy and this relates to the ability of a test to discriminate between patients who have a particular condition and those that do not (Šimundic, n.d.). A perfect diagnostic test is one that completely discriminates those with the disease and those without the disease, but this perfect test rarely exists (Simundic, n.d.).
3.3.1.1: Diagnostic accuracy

The discriminative potential of a diagnostic test can be quantified by measures of diagnostic accuracy. Which measures are used depends on whether the test is assessing discriminative properties or predictive ability (Šimundic, n.d.). Measures of diagnostic accuracy include sensitivity, specificity, predictive values, likelihood ratios, Youden’s index, diagnostic odds ratio and diagnostic accuracy.

Sensitivity is the probability that the test will produce true positive results when used on patients with the target condition, therefore recognising patients with the target condition (Programme for Appropriate Technology in Health (PATH) 2008). Specificity is the probability that the test will produce true negative results when used on non-affected patients, therefore recognising patients without the target condition (PATH 2008). Specificity compliments sensitivity and the two should always be reported together.

Predictive values are the probability a patient will or will not have the target condition (PATH 2008). The positive predictive value is the proportion of patients with a positive test in the total number of patients recording a positive test and having the target condition, and the negative predictive value is the proportion of patients with a negative test who do not have the target condition in the total number of patients with a negative result (Šimundic, n.d.).

Likelihood ratios are the ratio of expected test results in patients with the target condition to the subjects without the condition. A positive likelihood ratio tells us
how much more likely the positive test is to occur in subjects with the target condition compared to those without the condition, whereas negative likelihood ratios tell us the probability that the same result will occur in subjects without the target condition (Šimundic, n.d.). Both should be reported together. A positive likelihood ratio of $>10$ and a negative likelihood ratio of $<0.1$ indicate a good diagnostic test.

The diagnostic odds ratio is a global measure of accuracy of a test and is used for general estimation of discriminative power and comparison between two or more tests. It measures the ratio of the odds of positivity in patients with the target condition relative to the odds in patients without the target condition and is dependent on sensitivity and specificity (Šimundic, n.d.).

Youden’s Index is the oldest measure of diagnostic accuracy and is another global measure of accuracy and is used for the evaluation of overall discriminative power (Šimundic, n.d.). A score of ‘1’ indicates a perfect diagnostic test and a score of ‘0’ a poor test. Diagnostic accuracy or effectiveness is the proportion of correctly identified patients amongst all patients, and includes all the true positive and true negative results of a test.

The measures of diagnostic accuracy are not a fixed indication of a test's performance as some are affected by the prevalence of the target disease in the population under study and others by the spectrum and definition of the disease (Šimundic, n.d.). Because of this diagnostic test studies perform differently in different populations. For example, it is inappropriate to evaluate a
test in primary care that will ultimately be used in secondary care because the frequency and severity of the target condition will vary. The measures of diagnostic accuracy that are affected by disease prevalence and spectrum of disease can be found in table 8.

**Table 8: factors affecting measures of diagnostic accuracy**

<table>
<thead>
<tr>
<th>Measure of diagnostic accuracy</th>
<th>Factors affecting measures of diagnostic accuracy</th>
<th>Disease prevalence</th>
<th>Spectrum of disease</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity &amp; specificity</td>
<td></td>
<td>No</td>
<td>Yes</td>
<td>Results are transferable into other settings even with different prevalence of the target condition in the population</td>
</tr>
<tr>
<td>Predictive values</td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Should only be used when the test is used on a population that legitimately reflects the number of people with the target condition in the population. Not readily transferable into other study settings</td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td></td>
<td>No</td>
<td>Yes</td>
<td>Applicable only if definition of disease has not changed</td>
</tr>
<tr>
<td>Diagnostic odds ratio</td>
<td></td>
<td>No</td>
<td>Yes</td>
<td>Depends on criteria used to define target condition</td>
</tr>
<tr>
<td>Youden’s Index</td>
<td></td>
<td>No</td>
<td>Yes</td>
<td>Affected by the criteria used to define target condition</td>
</tr>
<tr>
<td>Diagnostic accuracy</td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Diagnostic accuracy increases as prevalence decreases but does not mean test is better in populations with low prevalence</td>
</tr>
</tbody>
</table>

In addition, the measures of diagnostic accuracy can also be affected by poor study design, which can limit the transferability of the results as levels of accuracy of the test can be under or overestimated (Šimundic, n.d.).

Phase one of this study comprises a diagnostic test study in which the accuracy of the tuning fork test in the detection of ankle fractures is assessed. Diagnostic accuracy within this test will be assessed by reporting sensitivity, specificity, positive and negative predictor values and diagnostic accuracy. Bachmann et al
(2003) identified that the pooled sensitivity of the OARs was 98% (CI 96-99) so it is anticipated that a sensitivity for the tuning fork test that is lower than the 95% CI interval of the OARs would not be acceptable to patients, clinicians, or health economists and as such sensitivity is the main measure for diagnostic accuracy of the tuning fork in phase one of the study.

3.3.1.2: Phase two – focus group discussions

Phase two of the study ran concurrently with the diagnostic test study undertaken in phase one of the study and included obtaining qualitative data collected from focus group discussions to obtain patient and clinicians experiences of the tuning fork test in clinical practice. Qualitative research is ‘a form of social enquiry that focuses on the way people make sense of their experiences and the world in which they live’ (Holloway & Wheeler 2010, p.3). It involves research that seeks to explore phenomena that enable researchers to understand the experiences, behaviours and feelings of the research participants. Whilst there are a number of different data collection methods used in qualitative research focus group discussions are defined by Sim (1998) as group interviews that focus on a specific topic with the aim of eliciting ideas, thoughts and perceptions of a lived experience.

The following chapter will focus on phase one of the study and will follow the Standards for Reporting Diagnostic Accuracy (STARD) checklist for diagnostic tests (table 9).
### Table 9: STARD checklist for reporting of studies of diagnostic accuracy (version January 2003)

<table>
<thead>
<tr>
<th>Section &amp; topic</th>
<th>Item #</th>
<th>On page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE/ABSTRACT/KEYWORDS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TITLE/ABSTRACT/KEYWORDS</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>METHODS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Participant sampling</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Data collection</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Test methods</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Test methods</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Test methods</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Test methods</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Test methods</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Statistical methods</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Statistical methods</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>RESULTS</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>RESULTS</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>RESULTS</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>RESULTS</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>RESULTS</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>RESULTS</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>RESULTS</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>RESULTS</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>RESULTS</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>RESULTS</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>RESULTS</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 4 – PHASE ONE: METHODOLOGY FOR ASSESSING THE DIAGNOSTIC ACCURACY OF THE TUNING FORK TEST IN THE DETECTION OF FRACTURES TO THE ANKLE

This chapter focuses on phase one of the study which incorporates a quantitative diagnostic test study to answer the following research question –

“Can the tuning fork test increase the diagnostic accuracy of the Ottawa ankle rules for patients attending emergency care settings following twisting ankle injuries?”

The following sections outline the study methodology and follow the STARD guidelines for reporting diagnostic studies (Bossuyt et al 2003).

4.1: Ethical & Regulatory Approval

This section outlines the ethical and regulatory approval required for the study and includes ethics, study monitoring, sponsorship and funding.

4.1.1: Ethics

Ethical approval was received from the Local NHS Research Ethics Committee, and the School of Health and Social Services ethics committee of the University of Portsmouth. Ethics is central to nursing
practice and all study participants had the right not to be harmed, the right to full disclosure, the right to self-determination and the right to privacy, anonymity and confidentiality as dictated by the Declaration of Helsinki 2008. The ethical principles of beneficence, non-maleficence, justice, autonomy, confidentiality and consent were relevant for this study and a summary of how each was addressed can be found in table 10.

4.1.2: Study Steering Group

A study steering group was established to monitor the progress of the study. The aim of the group was to ensure the study was carried out according to the protocol, to review the progress of the study, and to advise on the management of any significant issues that arose throughout the duration of the study. The group planned to meet every three months but due to adverse weather conditions over the winter months in 2009 and 2010 two meetings were cancelled. Group members included the designated principal investigator from each site, the research nurse, the study administrator, academic supervisors, the study statistician and I, as the chief investigator and chair of the group. Records of the meetings were kept by the study administrator. The minutes produced were checked by me before distribution to all steering group members.
**Table 10: Ethical principles relevant to the study**

<table>
<thead>
<tr>
<th><strong>Beneficence</strong> – Research should benefit the participant and society (Parahoo, 1997) and this study set out to reduce the number of x-rays requested in the future aiming to reduce the time a patient’s spends in the emergency care setting and reduce the exposure to radiation from unnecessary x-rays improving the quality of care for future patients. In addition, there would be potential cost savings to the NHS by reducing the number of x-rays requested.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-Maleficence</strong> - Research should do no harm and all the clinicians involved in the study were appropriately trained for their role. All received training on the use of the tuning fork test and how to complete the data collection forms. The test was non-invasive so it was estimated that potential harm from taking part in the study was minimal.</td>
</tr>
<tr>
<td><strong>Justice</strong> – Study participants had the right to fair and equitable treatment before, during and after their participation in the study (Polit &amp; Beck, 2004). Participants would continue to be assessed using the Ottawa ankle rules to determine the need for an ankle x-ray as per current practice.</td>
</tr>
<tr>
<td><strong>Autonomy</strong> - Participants were invited to volunteer to take part in the study after reading an information sheet. All participants were given as much time as they needed to decide whether to take part or not, and given the opportunity to ask questions. It was made clear that their decision would not affect the care they received during their ED visit.</td>
</tr>
<tr>
<td><strong>Confidentiality</strong> - Confidentiality and anonymity were respected at all times. All participants were given a study number and no personal detail was entered on any computer for study purposes. Participants are not identifiable in any report connected to the completion of the study. All primary study data was kept in a locked cupboard with access limited to myself, the research nurse and the study administrator.</td>
</tr>
<tr>
<td><strong>Consent</strong> - Potential participants were given a patient information sheet either at time of booking in by the receptionist or by the streaming/triage nurse in each department. When the departments were busy the information sheets were given to the patients by the assessing clinician. Prior to receiving the tuning fork test all participants were required to sign a consent form. Parents of a child 15 years and under were also required to give their written consent and the child assent.</td>
</tr>
</tbody>
</table>

### 4.1.3: Public Engagement

The study protocol, ethical and funding applications, and all study documentation were reviewed by a lay-member of the local public research ‘Engage’ group. This member, who was a free-lance journalist
and a retired health and safety officer, was also invited to attend the Study Steering Group meetings.

4.1.4: Sponsorship

The study was sponsored by an acute NHS trust in the South Central region.

4.1.5: Funding

The study received funding from the National Institute of Healthcare Research [NIHR] Central Commissioning Facility [CCF] Research for Patient Benefit programme (RfPB) to the sum of £211,000, payable over 2 years, and as such became a NIHR portfolio study. The funding enabled a full time research nurse and a part time administrator to be employed for the duration of the study to assist with the day to day running of the study across the three study sites. When the study expanded to include two new sites a request for additional funds was submitted to the NIHR and a further £9000 received. In addition, the study administrator’s contract was extended and funded for a further six months by the Comprehensive Local Research Network [CLRN]. As part of the funding contract 28 days notice had to be given for any publications about the study and the demographics of all recruited patients had to entered into the NIHR accruals database.
4.1.6: Edge™ database

In addition to the NIHR accruals database all patient demographics from three of the four study sites had to be entered onto the CLRN accrual database known as Edge™. This is a trials management database that allows research professionals and the CLRN to access a detailed ‘live’ database of on-going trials from any location (University of Southampton 2010). The fourth trust in which study site D was situated were not yet registered to use the Edge™ database.

4.2: Participants

The following sub-sections outline the study setting, sample size calculation, study population, inclusion and exclusion criteria and process of obtaining informed consent.

4.2.1: Study setting

The original study proposal included three study sites, all of which were located on the south coast of England. The study sites were to include a large urban ED, a nurse-led minor injury unit situated in a small market town, and an NHS walk-in-centre situated in a large city. However, the walk-in-centre withdrew from the study before recruiting any patients and two further mid-sized EDs from small towns in the south of England
joined the study in 2010. In order to maintain confidentiality the study sites have been coded A, B, C and D. A brief summary of each is given below.

- Study site A was a large urban ED of an acute NHS trust with a purpose built paediatric ED attached. This site was open 24-hours a day seven days a week and had an annual new attendance rate of circa 100,000. As well as being the Lead Investigator I was also the designated Principal Investigator at this site.

- Study site B was a nurse-led minor injury unit in a small market town with annual new attendances of circa 12,000. This unit is a satellite unit of study site A and is open from 8am to 9pm daily. The emergency nurse practitioners (ENP) working at this site also rotate to study site A. A senior ENP was designated the Principal Investigator at this site.

- Study site C is a mid-sized ED of an acute foundation NHS trust in a large market town with annual new attendance rates circa 55,000. This site is open 24-hours seven days a week. The ENP team leader was designated the Principal Investigator at this site.

- Study site D is a small ED of an acute NHS trust with annual new attendances of circa 40,000. It is open 24-hours a day seven days a week. The Principal Investigator at this site was a senior consultant in emergency care who had a special interest in minor injuries and sports injuries.
All the study sites had an established cohort of emergency nurse practitioners (ENPs) experienced in the assessment of minor injuries prior to the start of the study. In addition study sites A, B and C employed emergency care practitioners as part of their ENP service. Emergency care practitioners were nurses and paramedics who rotated between working in the ED and in the pre-hospital setting with the local ambulance trust. Study site A and C also employed consultant nurses who also autonomously assessed patients presenting to the departments. From herein these grades of staff will be grouped together under the term of ‘ENP’. Only study site A provided a 24-hour ENP service. All study sites had junior doctors rotating through them on two, three and six month placements, registrar trainees on six month and two-year placements, and consultants in emergency medicine. Herein this group of staff will be referred to as ‘doctors’.

4.2.2: Sample size calculations

An initial sample size calculation was performed using the Carley, Dosman, Jones and Harrison (2005) nomograms. This was calculated by assuming a prevalence of fractures in the population under study of 15%, a tuning fork sensitivity of 95% with a 95% CI. This revealed that a total of 180 patients were needed. However discussion with the study statistician revealed that this meant the study findings would be based on 20-30 fractures only and a decision to increase the number of fractures
in the study to circa 180 was made. The sample size was therefore recalculated to be 1300.

4.2.3: Study population

Potential participants were identified from consecutive patients as they presented to the study sites with an ankle injury following a simple twisting injury. Patients could ‘self-present’ to the study sites, be referred by another clinician, i.e. General Practitioner (GP), or be brought to the study site by ambulance. A minimum age limit of 12 years was set as the ENPs at the walk-in-centre included in the initial proposal were not permitted to request x-rays on children younger than 12 years of age. There was no upper age limit.

4.2.4: Inclusion & exclusion criteria

Inclusion and exclusion criteria were dictated by the suitability of the patient to be assessed by the OARs, the mechanism of injury and age of the patient. The patient must have been able to walk unaided prior to the injury and been assessed as OARs positive to the ‘malleolar zone’ as indicated by Stiell et al (1993) before being entered into the study. Inclusion and exclusion criteria can be found in table 11.
Table 11: Inclusion & exclusion criteria

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Age more than or equal to 12 years</td>
</tr>
<tr>
<td>• Attending the study site with a simple inversion or eversion ankle injury</td>
</tr>
<tr>
<td>• Ability to walk unaided prior to the injury</td>
</tr>
<tr>
<td>• No obvious deformity to the ankle</td>
</tr>
<tr>
<td>• Identified as having bony tenderness along the posterior aspect of the lateral and/or medial malleolus, and the distal fibula shaft as outlined in the OARs – otherwise known as the malleolar zone</td>
</tr>
<tr>
<td>• Ability to provide their own informed consent</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Age 11 years or less</td>
</tr>
<tr>
<td>• Obvious deformity to the ankle</td>
</tr>
<tr>
<td>• Alternative mechanism of injury other than inversion or eversion e.g. fall from height, road traffic accident</td>
</tr>
<tr>
<td>• Assessed as OARs negative i.e. no bony tenderness noted</td>
</tr>
<tr>
<td>• Tenderness to the foot zone only as identified by Stiell et al (1993)</td>
</tr>
<tr>
<td>• Non-traumatic ankle pain</td>
</tr>
<tr>
<td>• Diminished or altered sensation to the lower leg due to any mechanism / pathology</td>
</tr>
<tr>
<td>• Under the influence of drugs or drink</td>
</tr>
<tr>
<td>• Unable to walk prior to the injury (this is an exclusion for use of the OARs)</td>
</tr>
</tbody>
</table>

The main exclusion criteria were related to age of the patient, mechanism of injury, suitability for assessment with the OARs, and the patient’s ability to give informed consent. The patient must not have obvious deformity to the ankle or be deemed to be under the influence of alcohol or illicit drugs. Patients assessed as OARs negative or as having tenderness to the ‘foot zone’ only were not eligible to take part, as these patients do not require an ankle x-ray.
4.2.5: Informed consent

Potential participants aged 16 years and over were given a patient information sheet to read by either the reception staff or the triage nurse when they presented to one of the study sites. (Patient information sheet - appendix A). The information sheet included a summary of the background of the study, what taking part involved, what the patient could expect if they took part, the risks and potential benefits of taking part, and the details of the research team, funding and sponsorship. Potential participants aged 12-15 years were given an age appropriate information sheet, and their parents a parent information sheet containing the same information (Age appropriate and parent information sheet - appendix B).

After obtaining the history of the presenting complaint and undertaking a systematic clinical examination clinicians identified the patients who met the inclusion criteria and invited them to take part in the study. Patients were given time to read the information sheet and were given the opportunity to ask questions about the study before making their decision to participate or not. Patients were informed that the decision to take part in the study was completely voluntary, and that taking part in the study would not change the treatment they received for their injury. They were also informed that taking part meant they would still get the x-ray they required.
Once the patient had made a decision to take part they were asked to give their written consent by signing a consent form (Consent form - appendix C). If the participant was aged 12-15 years written consent was obtained from a parent and written assent from the child. At this stage of the study participants were also asked whether they consented to being contacted about a focus group discussion in preparation for phase two of the study. Patients were informed that consenting to take part in phase one of the study did not mean they were obliged to take part in the second phase of the study if they did not want to.

4.3: Index & reference test

All patients received the same index and reference test performed in the same sequence, index test first followed by reference test during the same visit to the study sites. The sub-sections below outline the reference test and the index test under investigation.

4.3.1: Index test: Tuning fork test

The index test under investigation was the tuning fork test on patients assessed as being OARs positive to the malleolar zone of their ankle. The tuning fork test comprises the clinician holding a flat ended Gardener Brown 128Hz tuning fork with the thumb and index finger of one hand whilst being struck against the palm of the other hand in order to make it vibrate without ringing (Dissmann & Han 2006). Without
touching the vibrating prongs the clinician places the flat end of the tuning fork against the patient’s ankle at the site of maximal tenderness, identified during clinical examination. The patient’s response is noted and the tuning fork re-struck against the hand and placed 6cm proximal to the first site, and the patient’s response again noted. Clinicians were informed that an increase in the sensation felt, or any verbal or non-verbal expression of discomfort or pain should be recorded as a positive tuning fork test.

Patients were not informed what the sensation of the tuning fork was like but did receive the tuning fork test to both ankles to ensure the patients knew what a ‘normal’ tuning fork test felt like on their ‘good’ ankle. Patients were selected as to which ankle was tested first by number sequencing of the data collection sheets. The data collection sheets were numbered ‘1’ and ‘2’. Number ‘1’ meant that the ‘good’ ankle was tested first and the number ‘2’ meant that the injured ankle was tested first (Data Collection Sheet - appendix D).

4.3.2: Reference test: Plain AP & lateral ankle x-rays

The reference test was the report of plain AP and lateral ankle x-rays, viewed electronically on the Picture Archiving Communications System (PACS) (Weatherburn, Bryan, Nicholas & Cocks 2000) by reporting radiographers and radiologist. Plain AP and lateral x-rays were chosen as they are routinely used within all the study sites to diagnose fractures.
as part of normal clinical practice. All patients received the reference test within three hours of the index test. The reference test was obtained by registered radiographers who were trained in musculoskeletal radiology and the x-rays interpreted by either reporting radiographers or radiologists experienced in the interpretation of musculoskeletal x-rays and the report entered onto the PACs system.

4.4: Training

No restriction was placed on the number of clinicians who could be involved in the study. Provided the clinicians had the skills to independently assess and manage minor injuries they were eligible to be involved in the study. This included nurses and doctors. All clinicians wanting to be involved in obtaining informed consent and performing the tuning fork test needed to attend training specific to the study before independently recruiting a patient into the study. I delivered three group training sessions during April and June 2009 for staff working at study sites A and B. The group training sessions included a summary of the study protocol, the process for obtaining consent, how to perform the tuning fork test and how to complete the data collection sheet. In addition, one-to-one training sessions were available throughout the duration of the study, by either the research nurse or I, either on request from the clinicians or when new clinicians joined the study site. Two group teaching sessions were held at each of the study sites C and D in April and May 2010 for all clinicians wanting to be involved in the study.
In addition the Principal Investigators at these each site received training on how to teach the tuning fork test technique to other clinicians. The research nurse visited all the study sites at least once weekly to deliver additional one-to-one training sessions as required.

Clinicians from all study sites were able to attend as many teaching sessions as they felt necessary and were signed as being competent at performing the tuning fork test by either the research nurse or myself. In addition all clinicians involved in the study were given a credit card sized laminated copy of the OARs and on the reverse a list of the inclusion and exclusion criteria as an aid memoir. Myself and the study research nurse held one-to-one training sessions for all new staff recruited at the sites throughout the duration of the study. All clinicians involved in recruiting patients into the study were experienced in the requesting and interpretation of ankle x-rays before the start of the study. All ENPs had undertaken an x-ray interpretation exam as part of their ENP training and all held the Ionising Radiation (Medical Exposure) Regulation certification (Great Britain Home Office 2000).

4.5: Statistical methods

The following sub-sections outline the statistical methods used and include data collection, data interpretation and data analysis techniques.
4.5.1: Data collection

The data collection sheet consisted of one side of A4 paper and was designed to be easy to complete by clinicians working in busy emergency care settings and who had obtained the patients consent, undertaken the initial clinical examination and performed the tuning fork test (Data collection sheet - appendix D). The majority of the data collection was undertaken prospectively by the assessing clinician. This included the date of attendance, age, gender and ethnicity of patient; the ankle injured (right or left); the degree of swelling and whether the bony prominence of the malleoli were visible; the presence of any distracting injuries or history of previous injury to the site; and the result of the tuning fork test and name of clinician performing the test. Table 12 summarises the data collected by the clinician at the time of patient assessment.

Table 12: Data collected by the assessing clinician

<table>
<thead>
<tr>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date of attendance</td>
</tr>
<tr>
<td>Age, gender and ethnicity of patient</td>
</tr>
<tr>
<td>Affected ankle (left or right)</td>
</tr>
<tr>
<td>OARs positive or inconclusive</td>
</tr>
<tr>
<td>Degree of swelling + ++ +++</td>
</tr>
<tr>
<td>Whether the bony prominences of the malleoli were visible or hidden by swelling</td>
</tr>
<tr>
<td>Previous injury to site</td>
</tr>
<tr>
<td>Presence of other injuries</td>
</tr>
<tr>
<td>The result of the tuning fork test on the injured and uninjured ankle.</td>
</tr>
<tr>
<td>Name of practitioner undertaking the tuning fork test</td>
</tr>
</tbody>
</table>
Clinical experience has identified that patients cannot always be classified as OARs positive or negative and in a number of patients the result of the OARs is unclear. These patients often go one to receive an ankle x-ray anyway and decision to include ‘Ottawa inconclusive’ on the data collection sheet was made.

4.5.2: Data interpretation & blinding

To exclude ‘test review’ or ‘diagnosis review bias’ (Sackett & Haynes 2002) the tuning fork test result for both ankles was recorded on the data collection sheet by the assessing clinician before the participant was sent for x-ray. Although the clinicians then had to interpret the x-ray in order to manage the injury the x-ray result used for the purpose of comparison with the index test was the one entered onto the PACs reporting system by a team of reporting radiographers and radiologists. All the x-rays were interpreted blind to the result of the tuning fork test by the reporting radiographers and radiologists. All the x-ray reports were then independently reviewed to assess whether the fractures met the classification of a clinically significant fracture set at the start of the study, which was all fractures except small avulsion fractures less than 3mm in depth. This review was performed blind to the tuning fork test result.

The data sets on the reverse of the data collection sheet were collected retrospectively by the research nurse within one week of the patient’s attendance and included the patient’s time of arrival and discharge
(taken from the patients notes), the time of x-ray (taken from PACs) and the x-ray report (transcribed from PACs). Each data sheet was colour coded to represent the different study sites in case of queries arising from missing, incomplete or illegible data. In addition, the research nurse reviewed the notes of the participants to ensure compliance with the inclusion and exclusion criteria. Where the research nurse was unsure whether the patient met the inclusion criteria or not I reviewed the notes and made the final decision. This was undertaken blind to the results of the tuning fork or x-ray result. All the data sets were given numerical values and input weekly into an Excel spread sheet by the research nurse.

4.5.3: Data monitoring group

A ‘virtual’ data monitoring group was established to assess the accuracy of the data input into the spread sheet. The data monitoring group consisted of the study statistician, a consultant in emergency medicine and an advanced ENP. It was their responsibility to independently check the data entered into the spread sheet for accuracy against the data collection sheets. A total of 20 out of every 100 participants were reviewed by the data monitoring group. Data found to be incorrect was brought to the attention of the research nurse and myself and amended as appropriate. At the end of the study I checked the complete data set for accuracy against all the data collection sheets. The data entered into the Excel spread sheet was then copied and pasted into the statistical
package SPSS 18 to aid data analysis. All categorical data were given numerical values to aid analysis.

### 4.5.4: Data analysis

Patient demographics were displayed in a table, showing exact numbers and percentages where appropriate. The mean and range of the age criterion were reported.

Sackett and Haynes (2002) recommend that the results of a diagnostic test study should report inconclusive and lost results as well as positive and negative results in order to assess the true accuracy of the two tests. The result of the tuning fork tests and the x-rays will initially be reported in a ‘3x3 table’ as shown in table 13. Results entered into cells v and z would indicate that the x-ray results were missing or indeterminate to fracture, cells w and y that the tuning fork test result was missing or indeterminate, and cell x missing or indeterminate x-ray and tuning fork test results.

<table>
<thead>
<tr>
<th>Table 13: 3x3 table</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray</td>
</tr>
<tr>
<td>Tuning fork test</td>
</tr>
<tr>
<td>Positive</td>
</tr>
<tr>
<td>Lost, not performed or indeterminate</td>
</tr>
<tr>
<td>Negative</td>
</tr>
</tbody>
</table>
Prior to the start of the study it was anticipated that any indeterminate or inconclusive x-ray or tuning fork test results would mean that the patient would be managed as a fracture until proven otherwise and therefore any data entered into these cells would be moved to cells b and c as appropriate. Moving the data from cells v, w, x, y, and z would enable a 2x2 table to be constructed from which sensitivity, specificity, positive and negative predictive values could be calculated using an on-line statistical calculator (Rosner 2006). Although the 3x3 table would include the total number of fractures identified in the study in the event of false negative tuning fork test results the 2x2 table would only include those fractures deemed to be clinically significant in the fracture column. Sensitivity, specificity, predictor values and diagnostic accuracy are reported as percentages (%) with 95% confidence intervals. It is anticipated that in order for the tuning fork test to be accepted by clinicians the sensitivity could not be lower than when the OARs are used in isolation, therefore sensitivity was deemed the most important measure of diagnostic accuracy.

Unadjusted and adjusted Odds Ratios (OR), calculated using a logistical regression model with SPSS 18.0, are reported to identify whether there were any predictor variables of a correct or positive tuning fork test. The variables of age, gender, ethnicity, previous or distracting injuries, swelling, ankle tested first, study site attended, and role of clinician performing the test were considered potential predictor variables. P values and 95% confidence intervals are reported. An OR above 1.0
indicates that the variable is predictive of a correct or a positive tuning fork test. P values equal to or less than 0.05 are considered clinically significant. After data analysis the study statistician checked all the results for accuracy.

4.6: Study website

A website was developed in collaboration with a local website designer (Tuning Fork 2011). The domain name was initially purchased and registered for two years and then extended by a further two years in 2011. Figure 15 shows a screen shot of the website home page.

![Figure 15: Screen shot of tuning fork study home page](image)
The website included the biographies of the research team, and summarised the aim, objectives and methodology of the study. In addition a forum enabled patients, clinicians and the public to contact the research team and leave feedback on their experiences of the tuning fork test and ask questions relating to the study process. A thermometer was used to give a pictorial representation of actual recruitment numbers. The research nurse was responsible for moderating the forum and the website was updated at regular intervals by the original designer. In addition, all clinicians involved in the study were given promotional pens throughout the duration of the study as an aide memoire for the website. The pens were printed with the logo ‘Think Ankle – Think tuning fork study’. The URL of the study website was also printed on the pens.

4.7: Findings

This section summarises the results of the diagnostic test study. After the initial section on recruitment and patient demographics it is divided into three sections

- The first looks at the performance of the tuning fork test as a diagnostic test for ankle fractures
- The second looks at the predictive variables that may affect the accuracy of the tuning fork test
- The third focuses on length of stay.
Chapter 4: Methodology for assessing the diagnostic accuracy of the tuning fork test in the detection of fractures to the ankle

4.7.1: Recruitment

Patients were recruited from June 1\textsuperscript{st} 2009 to June 30\textsuperscript{th} 2011 from study sites A and B, and from July 1\textsuperscript{st} 2010 to June 30\textsuperscript{th} 2011 at study sites C and D. Study site A and B are under the jurisdiction of the same acute hospitals NHS trust and during the study period a total of 8079 patients presented to these two sites with an ankle injury. 80% (6480) were diagnosed with a soft tissue injury and 78% (6309) received ankle x-rays. 42 patients were diagnosed with an ankle dislocation and 1405 with fractures to the ankle. This equates to fracture prevalence in those that were x-rayed of 22% across these two sites. Figure 16 summarises the number of ankle injuries attending study sites A and B throughout the duration of the study.

<table>
<thead>
<tr>
<th>Total ankle injury attendances at study A and B during study period</th>
<th>Number diagnosed with soft tissue injury / sprain</th>
</tr>
</thead>
<tbody>
<tr>
<td>8079</td>
<td>4862</td>
</tr>
<tr>
<td>Number sent home without x-ray</td>
<td>Number diagnosed with fracture</td>
</tr>
<tr>
<td>1770</td>
<td>1405</td>
</tr>
<tr>
<td>Number of ankle X-rays requested</td>
<td>Number diagnosed with ankle dislocation</td>
</tr>
<tr>
<td>6309</td>
<td>42</td>
</tr>
</tbody>
</table>

\textbf{Figure 16: Summary of ankle injuries attending study sites A & B during study period}
Chapter 4: Methodology for assessing the diagnostic accuracy of the tuning fork test in the detection of fractures to the ankle

The total number of ankle injuries presenting to study sites C and D throughout the duration of the study is unknown. This is because this data set was collected retrospectively after the study had finished and study sites C and D were unable to produce this information from their ED attendance records.

A total of 1358 patients were recruited into the study of which 45 were removed as they did not meet the inclusion criteria. Figure 17 summarises the number of patients recruited and removed from the study.

**Figure 17: Summary of patient recruited & removed from the study**

27 of the 45 were not simple twisting mechanism, but included road traffic accidents, falls from heights and direct trauma. Four of the 45 had obvious deformity to their ankle requiring manipulation in the ED, twelve
were reported as having tenderness to the foot only and received foot x-rays and not ankle x-rays, and two patients were below the age criterion at 8 and 11 years. There were no missing data sets and due to the nature of this phase of the study no patients were lost to follow-up.

When recruitment was plotted by month there were two peaks in recruitment, the first during August, September and October 2009, and the second in January, February and March 2011. The month of June and December were consistently low for recruitment. Recruitment by month is displayed in figure 18.

![Figure 18: Recruitment total by month](image)

1313 patients were included in the final analysis. Study site A recruited 63% (n = 830) of the participants, compared to 19% (n = 244), 15% (n =
193), and 4% (n = 46) for study sites B, C and D respectively. Figure 19 shows the total recruitment by each study site.

![Figure 19: Recruitment by study site](image)

**4.7.2: Patient demographics**

56% (n = 730) of the participants were male compared to 44% (n = 583) female. The mean age was 34 years with an age range of 12-91 years. The largest age category was the 21 to 30 year age group (n = 368). 54% (n = 703) of the population were aged less than 30 years. Only 8% (n = 103) were aged 61 years or over. Figure 20 shows the age distribution of the participants.
Table 14 summarises the demographic profile and presentation characteristics of the study participants. 98% (n = 1287) of the participants were of white ethnic origin, and the right ankle was injured in 51% (n = 673). 16% (n = 211) of patients reporting they had sustained an injury to the same ankle previously, and 9% (n = 117) of participants had an additional distracting injury at the time of presentation. All patients were screened for inclusion into the study using the OARs with only patients assessed as having tenderness to the malleolar zone being eligible for inclusion into the study however 10% (n = 126) of those recruited were reported as being OARs ‘inconclusive’ by the assessing clinicians. All these went on to receive ankle x-rays so were included in the final analysis.
Table 14: Demographic profile & presentation characteristics of the study participants (Figures are counts (%) unless otherwise stated)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study site</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>830</td>
<td>(63)</td>
</tr>
<tr>
<td>B</td>
<td>244</td>
<td>(19)</td>
</tr>
<tr>
<td>C</td>
<td>193</td>
<td>(15)</td>
</tr>
<tr>
<td>D</td>
<td>46</td>
<td>(4 )</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>12 - 91</td>
<td></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>730</td>
<td>(56)</td>
</tr>
<tr>
<td>Female</td>
<td>583</td>
<td>(44)</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1287</td>
<td>(98)</td>
</tr>
<tr>
<td>Asian</td>
<td>9</td>
<td>(&lt; 1)</td>
</tr>
<tr>
<td>Black</td>
<td>4</td>
<td>(&lt; 1)</td>
</tr>
<tr>
<td>Chinese</td>
<td>4</td>
<td>(&lt; 1)</td>
</tr>
<tr>
<td>Mixed race</td>
<td>7</td>
<td>(&lt; 1)</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>(&lt; 1)</td>
</tr>
<tr>
<td><strong>Ankle injured</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>673</td>
<td>(51)</td>
</tr>
<tr>
<td>Left</td>
<td>640</td>
<td>(49)</td>
</tr>
<tr>
<td><strong>Site of tenderness on examination</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral malleolus (LM)</td>
<td>903</td>
<td>(69)</td>
</tr>
<tr>
<td>Medial malleolus (MM)</td>
<td>84</td>
<td>(6 )</td>
</tr>
<tr>
<td>Distal fibula shaft (DFS)</td>
<td>44</td>
<td>(3)</td>
</tr>
<tr>
<td>LM &amp; MM</td>
<td>141</td>
<td>(11)</td>
</tr>
<tr>
<td>MM &amp; DFS</td>
<td>5</td>
<td>(&lt; 1)</td>
</tr>
<tr>
<td>LM &amp; DFS</td>
<td>92</td>
<td>(7)</td>
</tr>
<tr>
<td>LM, MM &amp; DFS</td>
<td>43</td>
<td>(3)</td>
</tr>
<tr>
<td>Not listed</td>
<td>1</td>
<td>(&lt; 1)</td>
</tr>
<tr>
<td><strong>OARs result</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Definite positive</td>
<td>1187</td>
<td>(90)</td>
</tr>
<tr>
<td>Inconclusive</td>
<td>126</td>
<td>(10)</td>
</tr>
<tr>
<td><strong>Degree of swelling</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>38</td>
<td>(3 )</td>
</tr>
<tr>
<td>+</td>
<td>481</td>
<td>(37)</td>
</tr>
<tr>
<td>++</td>
<td>642</td>
<td>(48)</td>
</tr>
<tr>
<td>+++</td>
<td>152</td>
<td>(12)</td>
</tr>
<tr>
<td><strong>Bony prominence visible</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>803</td>
<td>(61)</td>
</tr>
<tr>
<td>No</td>
<td>510</td>
<td>(39)</td>
</tr>
<tr>
<td><strong>Distracting injuries present</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>117</td>
<td>(9 )</td>
</tr>
<tr>
<td>No</td>
<td>1196</td>
<td>(91)</td>
</tr>
<tr>
<td><strong>Previous injury on ankle</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>211</td>
<td>(16)</td>
</tr>
<tr>
<td>No</td>
<td>1102</td>
<td>(84)</td>
</tr>
<tr>
<td><strong>Ankle tested first with tuning fork</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Injured ankle</td>
<td>690</td>
<td>(53)</td>
</tr>
<tr>
<td>‘Good’ ankle</td>
<td>623</td>
<td>(47)</td>
</tr>
</tbody>
</table>
60% (n = 794) of the participants were assessed as having significant swelling, indicated by ‘++’ or ‘+++’, whilst the bony prominence of the malleoli was still visible in 61% (n = 803). 3% (n = 38) were assessed as having no swelling at all.

69% (n = 903) had tenderness to the lateral malleolus, 6% (n = 84) the medial malleolus, and 3% (n = 44) to the distal fibula shaft. However, 11% (n = 141) were reported as having tenderness to both the lateral and the medial malleoli, 7% (n = 92) had tenderness to the lateral malleolus and the distal fibula shaft, and <1% (n = 5) had tenderness to the medial malleolus and the distal fibula shaft. In addition 3% (n = 43) were reported as having tenderness to all three areas of the ankle.

4.7.3: Performance of the tuning fork as a diagnostic test for ankle fracture

All 1313 participants included in the analysis received the tuning fork test to both ankles and then went on to have AP and lateral ankle x-rays. 210 participants had a ‘positive’ x-ray with a total of 232 fractures diagnosed. This equates to a prevalence of fractures in the population of 16%. The lateral malleolus was the bone most frequently injured, accounting for 97 (42%) of the fractures. There were 82 distal fibula shaft fractures accounting for 35% of the fractures in the population, and 31 medial malleolus fractures, accounting for 15% of the fractures. In addition,
there were fourteen talar fractures and eight posterior malleolus fractures diagnosed. Table 15 lists the distribution of fracture location.

**Table 15: Distribution of fracture location**

<table>
<thead>
<tr>
<th>Fracture site</th>
<th>N</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lateral malleolus</td>
<td>97</td>
<td>(42)</td>
</tr>
<tr>
<td>Distal fibula shaft</td>
<td>82</td>
<td>(35)</td>
</tr>
<tr>
<td>Medial malleolus</td>
<td>31</td>
<td>(13)</td>
</tr>
<tr>
<td>Talus</td>
<td>14</td>
<td>(6 )</td>
</tr>
<tr>
<td>Posterior malleolus</td>
<td>8</td>
<td>(3 )</td>
</tr>
</tbody>
</table>

There were a total of 611 negative, 663 positive and 39 inconclusive tuning fork test results. Table 16 summarises the diagnostic outcome for the tuning fork test and x-ray in a 3x3 table. All x-ray reports were reviewed independently and graded as to whether the fracture was deemed clinically significant or not.

**Table 16: Diagnosis outcome for fractures: Tuning fork versus x-ray**

<table>
<thead>
<tr>
<th>Tuning Fork Test</th>
<th>X - ray</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Positive</td>
<td>138</td>
</tr>
<tr>
<td>Positive</td>
<td>Inconclusive</td>
<td>0</td>
</tr>
<tr>
<td>Positive</td>
<td>Negative</td>
<td>525</td>
</tr>
<tr>
<td>Positive</td>
<td>Total</td>
<td>210</td>
</tr>
<tr>
<td>Inconclusive</td>
<td>Positive</td>
<td>0</td>
</tr>
<tr>
<td>Inconclusive</td>
<td>Total</td>
<td>0</td>
</tr>
<tr>
<td>Negative</td>
<td>Positive</td>
<td>7</td>
</tr>
<tr>
<td>Negative</td>
<td>Total</td>
<td>65</td>
</tr>
<tr>
<td>Negative</td>
<td>Positive</td>
<td>0</td>
</tr>
<tr>
<td>Negative</td>
<td>Total</td>
<td>65</td>
</tr>
<tr>
<td>Total</td>
<td>Positive</td>
<td>138</td>
</tr>
<tr>
<td>Total</td>
<td>Total</td>
<td>210</td>
</tr>
</tbody>
</table>
So although 65 fractures were missed by the tuning fork test 38 of these were deemed not to be being clinically significant, as defined at the start of the study, and were therefore moved into the x-ray negative column in the 2x2 table for analysis (as discussed in section 4.5.4 – Data analysis). In addition the 39 inconclusive tuning fork test results were moved to the tuning fork positive row of the table as these patients would all receive an x-ray if the tuning fork were introduced into clinical practice. Table 17 shows the adjusted results for analysis displayed in a 2x2 table.

**Table 17: 2x2 table showing relationship between tuning fork test & x-ray**

<table>
<thead>
<tr>
<th>Tuning Fork Test</th>
<th>X-ray</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Clinically significant fracture</td>
<td>Negative</td>
</tr>
<tr>
<td>Positive</td>
<td>145</td>
<td>557</td>
</tr>
<tr>
<td>Negative</td>
<td>27</td>
<td>584</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>172</td>
<td>1141</td>
</tr>
</tbody>
</table>

The tuning fork test was therefore accurate in 729 patients, with 145 true positive and 584 true negative results. Diagnostic accuracy was 56% (95% CI 53 - 58). Sensitivity is calculated as 84% (95% CI 78 – 89) and specificity 51% (95% CI 48 – 54). The percentage of patients correctly classified as negative by the tuning fork test (the Negative Predictive Value) is 96% (95% CI 94 - 97), and the percentage of patients correctly classified as positive by the tuning fork test (the Positive Predictive Value) is 21% (95% CI 18-24).
Based on these results x-rays could have been reduced from 1313 to 702, equating to a reduction of 47%, but this is at the expense of missing 27 clinically significant fractures. However, scrutiny of these 27 missed fractures revealed that eight were managed as soft tissue injuries. These include:

- five small avulsion fractures that were greater than 3mm but less than 4.5mm in depth
- an undisplaced buckle fracture to the distal fibula shaft
- a 5mm avulsion fracture to the lateral malleolus in which the patient presented 19 days post injury
- and a 17mm avulsion fracture to the lateral malleolus in a patient who also had a fracture to the 5th metatarsal and presented 3½ weeks post injury. This patient received fracture management for the metatarsal fracture but not the malleolar fracture.

The remaining clinically significant fractures missed by the tuning fork were:

- Ten oblique fractures to the distal fibula shaft
- Six avulsion fractures to the lateral malleolus, varying from 3.3mm to 8.6mm in depth
- One 25mm avulsion fracture to the lateral malleolus and
- Two bi-malleolar fractures - one involving the medial and posterior malleoli and one the lateral and posterior malleoli.

Table 18 summarises the false negative results.
Table 18: Summary of the 27 clinically significant fractures missed by the tuning fork

<table>
<thead>
<tr>
<th>Fracture diagnosed on x-ray for the tuning fork test false negative patients</th>
<th>Number</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avulsion fracture &gt;3mm but &lt;4.5mm</td>
<td>5</td>
<td>All managed as soft tissue injuries</td>
</tr>
<tr>
<td>5mm avulsion fracture to lateral malleolus</td>
<td>1</td>
<td>19 days post injury, managed as soft tissue injury</td>
</tr>
<tr>
<td>Buckle fracture to distal fibula</td>
<td>1</td>
<td>Managed as soft tissue injury</td>
</tr>
<tr>
<td>17 mm avulsion fracture to the lateral malleolus</td>
<td>1</td>
<td>3½ weeks post injury with fractured 5th metatarsal. Ankle injury managed as soft tissue injury</td>
</tr>
<tr>
<td>Oblique fracture to distal fibula shaft</td>
<td>10</td>
<td>1 reported as having tenderness over medial malleolus 7 reported as having tenderness over lateral malleolus</td>
</tr>
<tr>
<td>Avulsion fracture 3.3mm to 8.6mm</td>
<td>6</td>
<td>Managed in plaster cast</td>
</tr>
<tr>
<td>25mm avulsion fracture to the lateral malleolus</td>
<td>1</td>
<td>Tender medially</td>
</tr>
<tr>
<td>Bi-malleolar fractures</td>
<td>2</td>
<td>Managed in plaster cast</td>
</tr>
</tbody>
</table>

When the ED records of these patients were reviewed the patient with the 25mm avulsion fracture to the lateral malleolus and one patient with a fracture to the distal fibula shaft were recorded as having tenderness to the medial aspect of the ankle rather than laterally. In addition, seven of the patients diagnosed with fractures to the distal fibula shaft were recorded as having bony tenderness to the tip of the lateral malleolus and not the distal fibula shaft.

4.7.4: Predictor variables influencing the accuracy of the tuning fork test

A total of 113 clinicians were involved in performing the tuning fork test across the four study sites - 60 non-medical clinicians consisting of
emergency nurse practitioners, emergency care practitioners and consultant nurses, and 53 doctors of variable experience from senior house officer to consultants in emergency medicine. This equates to a mean of 12 participants (range 1 to 127) per clinician. The non-medical clinicians entered 1079 (82%) of patients into the study, equating to a mean of 18 per nurse, compared to 234 (18%) entered by doctors, equating to a mean of only 5 per doctor. However, this number skewed by the fact that one ENP recruited 127 patients and one doctor 89. Removing these two clinicians from the analysis and the nurses recruited a mean of 15 patients each and the doctors 3.

4.7.4.1: Correct tuning fork test

Based upon the information recorded at presentation, it was possible to investigate what factors, if any, might be associated with a correct tuning fork test. Ethnicity was removed from the analysis due to the low number of patients from ethnic groups other than white but the variables of age, gender, degree of swelling, distracting or previous injury to the site, study site attended and role of clinician performing the tuning fork test were all considered potential confounders to an accurate tuning fork test. Table 19 records the outcome of a logistic regression analysis in which each predictor variable is first considered individually (unadjusted analysis) and secondly when included in a multiple logistic regression model including all possible predictors (adjusted analysis) for diagnostic
accuracy. Marked differences between the unadjusted and adjusted OR is considered an indication of confounding.

**Table 19: Variables predictive of an accurate tuning fork test**

<table>
<thead>
<tr>
<th>Predictor variable</th>
<th>Unadjusted (Crude) analysis</th>
<th>Adjusted analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds Ratio (CI)</td>
<td>p value</td>
</tr>
<tr>
<td>Patient age</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.020 (1.013-1.028)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Gender: Male</td>
<td>1.116 (0.897-1.389)</td>
<td>0.325</td>
</tr>
<tr>
<td>Female</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Injured Ankle:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>0.968 (0.779-1.203)</td>
<td>0.770</td>
</tr>
<tr>
<td>Left</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Degree of swelling (overall)</td>
<td>1</td>
<td>1.275 (0.658-2.470)</td>
</tr>
<tr>
<td>None</td>
<td>1.349 (0.700-2.598)</td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>2.137 (1.041-4.388)</td>
<td></td>
</tr>
<tr>
<td>++</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>+++</td>
<td>1.232 (0.985-1.542)</td>
<td>0.068</td>
</tr>
<tr>
<td>Visible bony prominence</td>
<td>1</td>
<td>1.070 (0.731-1.566)</td>
</tr>
<tr>
<td>Distracting injuries present</td>
<td>1</td>
<td>0.974 (0.724-1.310)</td>
</tr>
<tr>
<td>Previous injury to ankle</td>
<td>1</td>
<td>0.842 (0.677-1.047)</td>
</tr>
<tr>
<td>“Good” ankle tested first</td>
<td>1</td>
<td>0.979 (0.698-1.327)</td>
</tr>
<tr>
<td>Study site (overall)</td>
<td>1</td>
<td>1.170 (0.852-1.607)</td>
</tr>
<tr>
<td>A</td>
<td>1.712 (0.910-3.218)</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.929 (0.698-1.237)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1.170 (0.852-1.607)</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>1.712 (0.910-3.218)</td>
<td></td>
</tr>
<tr>
<td>Test performed by:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENP</td>
<td>1.637 (1.232-2.176)</td>
<td>0.001*</td>
</tr>
<tr>
<td>Doctor</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Key: *statistically significant, p less than 0.05

Individual analysis (unadjusted) revealed that patient age (p <0.001), degree of swelling (p 0.039) and role of person performing the tuning fork
Chapter 4: Methodology for assessing the diagnostic accuracy of the tuning fork test in the detection of fractures to the ankle

Test (p 0.001) were significant predictors of an accurate tuning fork test with visibility of the bony prominence of the malleoli borderline (p 0.068). The variables of gender (p 0.325), ankle injured (p 0.770), presence of distracting injuries (p 0.728), previous injury (p 0.860), ankle tested first (p 0.121), and study site visited (p 0.242) were not significant predictors of an accurate tuning fork test.

After adjustment for potential confounding age (p <0.001) and role of clinician performing the test (p 0.003) were found to be significantly associated with outcome. Gender was borderline (p 0.078). The adjusted OR for age, 1.020, is more than 1.0 indicating that a correct result is more likely with increasing patient age. Likewise, the adjusted OR for male gender, 1.247, indicates that males are more likely to produce an accurate result than females, and finally, the adjusted OR for clinician role, 1.595, indicates that nurses are more likely to record a correct result compared to doctors. This is confirmed in table 20 below which shows that nurses recorded an accurate tuning fork test in 623 out of 1079 (58%) patients they assessed compared to 106 out of 234 (45%) for the doctors.

Table 20: Accuracy of the tuning fork by clinician role: nurses v doctors

<table>
<thead>
<tr>
<th>Role of clinician</th>
<th>Accurate tuning fork test</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Nurse</td>
<td>456</td>
<td>623</td>
</tr>
<tr>
<td>Doctor</td>
<td>128</td>
<td>106</td>
</tr>
<tr>
<td>Total</td>
<td>584</td>
<td>729</td>
</tr>
</tbody>
</table>
The degree of swelling was clinically significant (p 0.039) when single predictor analysis was performed but not when adjusted for confounders (p 0.175). Table 21 shows how the degree of swelling affected the accuracy of the tuning fork test. The tuning fork test appears to be more accurate in the presence of swelling as the percentage of those recording an accurate tuning fork test increases with the degree of swelling from 46% with no swelling to 66% with significant swelling (+++).

Table 21: Accuracy of the tuning fork in the presence of swelling

<table>
<thead>
<tr>
<th>Degree of swelling</th>
<th>Accurate tuning fork test</th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>(%)</td>
<td>Yes</td>
<td>(%)</td>
</tr>
<tr>
<td>None</td>
<td>20 (53)</td>
<td>18 (47)</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>223 (46)</td>
<td>258 (54)</td>
<td>481</td>
<td></td>
</tr>
<tr>
<td>++</td>
<td>290 (45)</td>
<td>352 (55)</td>
<td>642</td>
<td></td>
</tr>
<tr>
<td>+++</td>
<td>52 (34)</td>
<td>100 (66)</td>
<td>152</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>585</td>
<td>728</td>
<td>1313</td>
<td></td>
</tr>
</tbody>
</table>

There were no marked differences in the OR for the adjusted and unadjusted analysis for any variable so no confounders to an accurate tuning fork were identified.
4.7.4.2: Positive tuning fork test

In addition to analysing for predictors of a correct tuning fork a second analysis was performed to identify if any of the same variables were predictors of a positive tuning fork test, whether this was an accurate result or not. Table 22 records the outcome of a logistic regression analysis in which each predictor variable is again considered individually (unadjusted analysis) and in a multiple logistic regression model including all possible predictors (adjusted analysis) for a positive tuning fork test. Marked differences between the unadjusted and adjusted OR is considered an indication of confounding.

Based on the individual analysis the role of the clinician performing the test (p <0.001) and study site (p 0.045) were identified as being significantly associated with a positive test with male gender (p 0.078) borderline. Whereas the variables for age (p 0.102), ankle injured (0.685), degree of swelling (p 0.261), visible bony prominence (p 0.699), distracting injuries (p 0.515), previous injury (p 0.454), and ankle tested first (p 0.296) were not predictive of a positive test. After adjustment for potential confounding age (p 0.041), gender (p 0.016) and clinician performing the test (p <0.001) were found to be significantly associated with outcome.
Chapter 4: Methodology for assessing the diagnostic accuracy of the tuning fork test in the detection of fractures to the ankle

Table 22: Factors predictive of a "positive" tuning fork test

<table>
<thead>
<tr>
<th>Predictor variable</th>
<th>Unadjusted (Crude) analysis</th>
<th>Adjusted analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds Ratio (CI)</td>
<td>p value</td>
</tr>
<tr>
<td>Age</td>
<td>0.994 (0.998 - 1.001)</td>
<td>0.102</td>
</tr>
<tr>
<td>Gender: Male</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.821 (0.660 - 1.022)</td>
<td>0.078</td>
</tr>
<tr>
<td>Injured Ankle:</td>
<td>1</td>
<td>0.685</td>
</tr>
<tr>
<td>Right</td>
<td>1.046 (0.842 - 1.299)</td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree of swelling (overall)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1</td>
<td>0.999 (0.515-1.935)</td>
</tr>
<tr>
<td>+</td>
<td>0.994 (0.516-1.915)</td>
<td></td>
</tr>
<tr>
<td>+++</td>
<td>1.419 (0.694-2.901)</td>
<td></td>
</tr>
<tr>
<td>Visible bony prominence</td>
<td>1.045 (0.836 - 1.305)</td>
<td>0.699</td>
</tr>
<tr>
<td>Distracting injuries present</td>
<td>0.881 (0.601 - 1.291)</td>
<td>0.515</td>
</tr>
<tr>
<td>Previous injury to ankle</td>
<td>1.119 (0.833 - 1.503)</td>
<td>0.454</td>
</tr>
<tr>
<td>“Good” ankle tested first</td>
<td>1.123 (0.904-1.396)</td>
<td>0.296</td>
</tr>
<tr>
<td>Study site (overall)</td>
<td>1</td>
<td>0.045*</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.928 (0.697-1.236)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.760 (0.555-1.040)</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>0.469 (0.254-0.897)</td>
<td></td>
</tr>
<tr>
<td>Test performed by:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENPs</td>
<td>0.491 (0.364-0.662)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Doctor</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Key: *statistically significant, p less than 0.05

The adjusted OR for age, 0.993, is less than 1.0 indicating that a positive tuning fork test was less likely to occur with increasing patient age.

Likewise, the adjusted OR for male gender, 0.753, indicate that males...
were less likely to produce a positive tuning fork test compared to females, and finally, the adjusted OR for clinician role, 0.485, indicates that ENPs were less likely to declare the test positive compared to doctors which can be seen in table 23.

<table>
<thead>
<tr>
<th>Role of clinician</th>
<th>Tuning fork test result</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>ENPs</td>
<td>534</td>
<td>545</td>
</tr>
<tr>
<td>Doctor</td>
<td>76</td>
<td>158</td>
</tr>
<tr>
<td>Total</td>
<td>610</td>
<td>703</td>
</tr>
</tbody>
</table>

ENPs recorded 545 out of 1079 (51%) patients they recruited as positive, compared to 158 out of 234 (68%) for the doctors.

There were no marked differences in the adjusted and unadjusted odds ratios for any variable so no confounders were identified.

4.7.5: Time outcomes

An additional outcome of the study was the time the patients waited to be seen by the assessing clinician and the total length of stay from time of arrival.
The mean time from arrival to being seen by a clinician was 49 minutes (range 0 – 240 minutes). The mean time from assessment by a clinician to discharge was 62 minutes. When only the data from the patients recording a negative tuning fork test were reviewed the mean time from assessment to discharge was 66 minutes (Standard Deviation 43.3) therefore sending patients who recorded a negative tuning fork test to x-ray in this study extended their stay by more than an hour.

The mean total length of stay, from arrival to discharge, was 109 minutes, with 63% of patients being discharged within 2 hours, 89% within 3 hours and 99% within 4 hours of arrival.

### 4.8: Discussion of findings

This section summarises the findings of the diagnostic test study, including recruitment, population, operator experience, and accuracy. It will also outline the strengths and weaknesses of this phase of the study.

### 4.8.1: Recruitment

The recruitment target was 1300 and a total of 1313 data sets were available for analysis. This is therefore the largest study to report on the accuracy of the tuning fork test to date. Not surprisingly study site A recruited the most participants by entering 830 into the study. This equates to 63% of the total study population. Study site A is a large
urban emergency department which also includes a dedicated paediatric emergency department, and has a combined annual new patient attendance rate of approximately 100,000. This is in comparison to study sites B, C and D whose average new annual patient attendance rates are circa 12,000, 55,000 and 40,000 respectively. At study sites B, C and D recruitment was 19% (n=244), 15% (n=193) and 4% (n=46) of the study population respectively. Study site B recruited the larger proportion of their annual attendance figures when compared to the other study sites. This was possibly because study site B was a dedicated minor injury unit and as such ankle injuries would have made up a larger proportion of their attendances.

4.8.2: Study population

The majority of studies on the incidence of ankle injuries focus on young males who have injured their ankle during sport and 56% of the population in this study were male. The majority of the population were aged less than 40 years of age (70%, N = 921). Patients had to have sustained a twisting ankle injury but the data collection did not require the exact mechanism to be recorded. On reflection this data would have been beneficial to the study. This data could have added to the knowledge on the common mechanisms of injury that cause twisting ankle injuries, and would have identified the number of participants that injured their ankle during sport, at home, at work and during other recreational activities.
1287 (98%) of the study population were from a white ethnic background with the remainder made up of Asian (n = 9), Black race (n = 4), Chinese (n = 4), Mixed race (n = 7) and other (n = 2). Unfortunately the withdrawal of the walk-in-centre affected the mix of ethnicity expected in the population under study. The walk-in-centre was located in a densely populated urban city, in an area with a high multi ethnic population and following their withdrawal only 26 study participants were classified as being from a non-white ethnic group. This variable therefore had to be excluded from analysis due to the low numbers included and limits the transferability of the results into these patient groups.

Garrick (1977) states that the lateral side of the ankle is the most common site of injury and 93% (n = 1223) of patients in this study presented with tenderness to the lateral aspect of their ankle compared to only 6% (n = 84) with isolated medial tenderness. 210 participants were diagnosed with a fracture on x-ray, which equates to a prevalence of fractures in the population under study of 16% and reflects the prevalence seen nationally (Struijs & Kerkhoffs 2010). Lateral malleolus fractures accounted for 42% (n = 97) of fractures and distal fibula shaft fractures for 35% (n = 82). The OARs does not include palpation of the posterior malleolus or the talus but eight patients were diagnosed with fractures to the posterior malleolus and 14 to the talus, in total accounting to 6% of the fractures in this study. None of the posterior malleolus fractures were isolated injuries, all occurring in association with lateral or medial malleolar fractures.
4.8.3: Operator experience

It was important that the accuracy of the tuning fork was studied in the clinical setting to which it would be used if implemented and as such this study did not restrict the number of clinicians involved, unlike the Dissmann and Han (2006) study which restricted the use of the tuning fork to a single operator. All clinicians were eligible to recruit patients provided they autonomously assessed patients as part of their routine clinical work and attended the training sessions prior to recruiting patients into the study.

A total of 113 clinicians (60 nurses and 53 doctors) were involved in recruiting patients, obtaining consent, performing the tuning fork test and collecting study data over the duration of the study, making this the largest number of clinicians involved in a study that assesses the use of the tuning fork in the detection of fractures. The study sites employed nurses working in a variety of roles, such as ENPs and consultant nurses, to manage minor injuries prior to the start of the study and it was not surprising that as a group the nurses recruited 82% \((n = 1079)\) of the patients enrolled into the study. This is compared to 18% \((n = 234)\) for the doctors.

It became evident at the training sessions that all of the nurses had a prior working knowledge of the OARs and the training focused on ensuring they were all using the same version of the OARs and that they
knew how to apply and interpret the tuning fork test, whereas the junior doctors appeared to have no knowledge of the OARs prior to joining the study. However, they all received training on the use of the OARs as well as how to apply and interpret the tuning fork test at the training sessions.

When clinician role was analysed individually and in a multiple logistical regression model operator role was found to be a predictor of an accurate tuning fork test. With an adjusted OR of 1.595 (95% CI 1.174-2.167) nurses were more likely to record an accurate result than doctors (p 0.003). Doctors also recorded a higher proportion of positive tuning fork tests when compared to nurses (68% and 51% respectively). Since nurses were frequently assessing ankle injuries, often daily, they may have become proficient in their assessment. Nurses also recruited more patients into the study and therefore were more familiar with using and interpreting the tuning fork test than the doctors.

4.8.4: Inconclusive results

This is the first study on the OARs to report inconclusive results. 126 (10%) of the study population were recorded as OARs inconclusive. All these patients went on to receive an x-ray and were therefore included in the final analysis. The study did not set out to establish what clinicians determined was an inconclusive result but 10% of the total study population is a significant amount and warrants further investigation. In addition, 39 (3%) patients recorded an inconclusive tuning fork test and it
was assumed, for the purpose of the study, that these patients would go on to receive an x-ray if the test was introduced into clinical practice. Again, the study did not set out to establish what an inconclusive tuning fork test was and further research into this is recommended.

4.8.5: Missed fractures

Colwill and Berg (1958), Bache and Cross (1984), Misurya et al (1987), and Moore (2009) all report false negative tuning fork tests in their papers and there were 65 false negatives in this study. However, 38 of the 65 reported in this study were deemed not to be clinically significant (Stiell et al 1992). Review of the attendance records of these patients confirmed they had all been managed as soft tissue injuries so these were transferred from the ‘x-ray positive’ column of the 2x2 table to the ‘x-ray negative’ column for analysis. This meant a total of 27 clinically significant fractures were missed by the tuning fork test. However, eight were further managed as soft tissue injuries and given no follow-up; five were small avulsion fractures measuring between 3mm and 4.5mm and one was an undisplaced buckle fracture to the distal fibula and two larger avulsion fractures to the lateral malleolus in patients that presented 19 days and 3 ½ weeks post injury. This indicates that there may be a time limit on when the tuning fork test is accurate in the detection of an acute fracture. This is in contrast to the detection of stress fractures in which it is stated the tuning fork is more accurate in the advanced stages of the fracture process (Lesho 1997).
Dissmann and Han (2006) report that the tuning fork test was more accurate when applied to the distal fibula shaft than the lateral malleolus but throughout the duration of the study it became apparent to clinicians that the tuning fork test was missing distal fibula shaft fractures. Ten of the false negative tuning fork tests were fractures to the distal fibula shaft. However, in seven of these the patients were recorded as having maximum tenderness to the tip of the lateral malleolus and one as having tenderness to the medial malleolus. In addition, a patient with a large avulsion fracture to the lateral malleolus, missed by the tuning fork test, was also reported as having tenderness over the medial malleolus.

This could, therefore, imply that in nine patients, recorded as false negative, the tuning fork could have been incorrectly placed due to either clinician error in judgement of the site of tenderness or the fact that in these injuries patients have associated significant soft tissue injury that distract the pain away from the actual fracture site. Unfortunately data analysis did not take place to assess individual clinician accuracy in interpreting the tuning fork test due to the wide variation in numbers of patients seen by individual clinicians, which ranged from 1 to 127.

4.8.6: Accuracy of the tuning fork test

A total of 663 (50%) patients recorded a positive and 611 (47%) a negative tuning fork test result. The tuning fork test was correct in 729, equating to a diagnostic accuracy of 56% (95% CI 53-58). Despite a high
NPV (96%, 95% CI 94-97) and higher specificity (51%, 95% CI 48-54) than when the OARs are used in isolation (39%, Interquartile range 28-48) sensitivity of the tuning fork test is reported as 84% (95% CI 78-89) which is lower than when the OARs are used in isolation (Bachmann et al 2006). It is therefore unlikely that this will be acceptable to patients or clinicians as this equates to approximately one in six fractures being missed by the tuning fork test. The CI for sensitivity, specificity and NPV are narrow meaning that the study had sufficient power to detect effect.

4.8.7: Predictor variables of a correct & positive tuning fork test

Although the main aim of this phase of the study was to assess the diagnostic accuracy of the tuning fork test a second outcome was to identify whether there were any predictor variables for an accurate or a positive tuning fork test, and this is the first study on the use of a tuning fork to detect fractures to assess this.

The variables of age, gender, swelling, distracting or previous injury, ankle tested first, study site attended, and role of the clinician performing the test were all considered potential confounding variables to an accurate and a positive test. As previously stated ethnicity had to be excluded from the analysis as the study population did not include sufficient patients from ethnic backgrounds. Although the majority of the variables were not identified as predictors to an accurate or a positive
tuning fork test patient age and role of the clinician performing the test were. Swelling was a predictor variable when analysed individually only.

Unlike Dissmann and Han (2006) this study included patients with swelling. This would ensure that the study population was a true reflection of the patient group the tuning fork test would be applicable to if introduced into clinical practice. 97% (1275) of the study population were recorded as having some degree of swelling, whilst only 3% (38) had no swelling. Clinicians regularly use a ‘+’ symbol to indicate the degree of swelling, with ‘+’ indicating minimal swelling and ‘+++’ to indicate significant swelling. As these symbols were widely used across all the study sites before the start of the study clinicians were not instructed as to what the symbols meant on the data collection sheet, and this was left to individual interpretation. In total 60% (794) of patients were recorded as having ‘++’ or ‘+++’ of swelling yet the malleoli was recorded as still being visible in 61% (803) and I would argue that in the presence of significant swelling the malleoli would not be visible.

Dissmann and Han (2006) anticipated that swelling would affect the result of the tuning fork test and with an OR of 2.137 (95% CI 1.041-4.388) for ‘+++’ of swelling and an OR of 1.0 for no swelling an accurate tuning fork test was actually more likely to occur in the event of significant swelling (p. 0.039), which is the reverse of what Dissmann and Han (2006) anticipated. However, there were four times as many patients recorded as having significant swelling (152) than no swelling (38) so this result needs to be treated with caution.
The variable ‘patient age’ shows that a positive tuning fork test is less likely with increasing age (adjusted OR 1.021, 95% CI 1.014-1.029, p. <0.001) and an accurate test more likely to occur with increasing age (adjusted OR 0.993, 95% CI 0.985-1.000, p 0.041). Both the result for age and role of clinician are clinically significant as the p value is less than 0.05.

4.8.8: Length of stay

The final outcome of this study was to assess how a patient’s length of stay could be affected if the tuning fork test were introduced into clinical practice. The majority of participants (63%) were discharged within two hours and 99% within four hours of arrival meeting the DH target for total length of stay in the ED (DH 2000). It is stated that sending a patient to x-rays adds more than 30 minutes to their ED stay (Anis et al 1995; Clements 2004 cited in Dissmann & Han 2006) but the mean time from assessment by the clinician to time of discharge in this study was 62 minutes. If the patients recording a negative tuning fork test had been discharged immediately after assessment, their length of stay in the ED would have been reduced by 66 minutes which is double the time stated by Anis et al (1995) and Clements (2004, cited in Dissmann & Han 2006).
4.9: Strengths & weaknesses of this phase of the study

In order that the validity of a diagnostic test study can be assessed, Devillé and Buntinx (2002) state that internal and external validity has to be established. Internal validity is affected by whether the patients are recruited consecutively or retrospectively, which reference and index test is used, how and when these are applied, and whether the results are interpreted blind (Begg 1987, cited in Irwig et al 2002). External validity is affected by the population under study, the study setting, patient demographics, duration of the condition before diagnosis, the presence of co-morbid conditions, and missing data sets (Devillé & Buntinx 2002).

The aim is to avoid introducing biases into the study that can invalidate or reduce the generalizability of the findings (American Medical Association n.d.; Whiting et al 2003). How each of these has been addressed in this study is discussed below.

4.9.1: Consecutive patients

Validity of any diagnostic test study is affected by whether patients are recruited consecutively and whether data is collected retrospectively or prospectively (Devillé & Buntinx 2002). The study process set out that all patients be consecutively recruited into the study. However, it soon became apparent that at times when the ED was busy or the clinicians were working in isolation they appeared reluctant to recruit into the study possibly due to the additional work involved adding to their already busy
workload. This may have comprised consecutive recruitment of patients. However, no patients were entered into the study retrospectively.

The majority of the data was collected prospectively by the assessing clinician to improve validity, but the x-ray report and the time data sets were collected retrospectively by the research nurse. This took account of the time lag between the x-ray being taken and the report being made available on PACs.

4.9.2: Training of clinicians

All clinicians involved received training on the use of the tuning fork test and their knowledge of the OARs was reviewed. All the ENPs had a good working knowledge of the OARs but the junior doctors did not, so they were also instructed on the use of the OARs. After training clinicians were left to assess the patient with the OARs and perform the tuning fork test independently. No further assessment into the use of the tuning fork was made unless requested by the individual clinician. Only one ENP accessed additional training, but it became apparent at the focus group discussions that not all ENPs were convinced they were ‘doing things right’. The tuning fork test was performed by a single operator on each patient so it was not possible to check the accuracy of the interpretation of the results and in nine of the fractures missed by the tuning fork there is evidence that the tuning fork may have been incorrectly placed in
these patients. In future studies it is recommended that two operators perform the test on each patient and then analysis of inter-operator reliability can be performed.

4.9.3: Index and reference tests

Another test of validity is whether an adequate reference test has been used. In a diagnostic test the test under review needs to be compared to a ‘gold-standard’ reference test. According to Simon (1997) this reference test should be a slower, less convenient or more expensive route to giving a definitive diagnosis. The radiographer / radiologist report of plain AP and lateral ankle x-rays were chosen as an appropriate reference test in this study. Although a reference test should provide 100% sensitivity and specificity Knottnerus and Van Weel (2002) warn that this gold-standard test rarely exists, and it is known that x-rays do not always demonstrate a fracture (Hendrix 1992). This may be due to underexposure, insufficient views, patient movement and patient size (Hendrix 1992, cited in Van den Berg 2003). Fractures can also be very subtle and not always easily identifiable on x-ray. In addition, interpretation of x-rays relies on subjective assessment from human observers with varying levels of skill and this can also affect the accuracy of the interpretation (Weinstein, Obuchowski & Lieber 2005). However, all the study sites used plain x-rays to detect fractures in the acute setting and they were therefore deemed the most appropriate reference test to use.
Differential verification bias can be introduced into a diagnostic study if the patients receive different screening and reference tests (Devillé & Buntinx 2002; American Medical Association n.d.). All the patients in this study were screened using the OARs and all those recruited received the tuning fork test before being sent for x-ray. The description of the tuning fork test was the same as that used by Dissmann and Han (2006) to ensure a comparison could be made. Previous research into the use of the tuning fork to detect fractures used a 128Hz and therefore the same was used throughout this study. All the tuning forks had been purchased from the same supplier for the purpose of the study to ensure consistency.

Whiting et al (2003) also state that disease progression bias can be introduced if the two tests are not performed within a reasonable time frame due to progression of the disease over time. It can be argued that Lesho (1997) introduced disease progression bias into his study on the use of the tuning fork to detect stress fractures as the tests were performed 30 days apart. Within this study disease progression bias was eliminated as the tuning fork test and the x-ray were performed in the majority of patients within one hour of each other, and in all instances within three hours.
4.9.4: Blinding

Blinding of the test results is an integral part of ensuring that the results of a diagnostic test are credible and valid and eliminates test review and diagnostic review bias from a study (Kelly et al 1997). There is a risk that diagnostic review bias may have been introduced into this study as the clinicians performing the tuning fork test had to interpret the x-ray result in order to manage the patient’s injury correctly. However, the patient’s pathway through the study meant the result of the tuning fork test had to be recorded before the patient went to x-ray, and the x-ray report used for comparison was that entered onto PACs by the radiologist or reporting radiographer and not that entered by clinicians in the patients records. Therefore, it is anticipated that diagnostic review bias is minimal. Test review bias was completely eliminated as none of the radiologists or reporting radiographers knew which patients were recruited into the study and the report of their x-rays was undertaken as part of routine reporting of all x-rays at the study sites.

4.9.5: Sample size

Diagnostic test studies are traditionally small scale and tend not to report sample size calculations, despite the recommendation in the STARD checklist (Bossuyt et al 2003). This is evident in the meta-analysis by Bachmann et al (2006) who identified that of the 43 papers reviewed for their meta-analysis only 2 included sample size calculations. In order to
overcome this, the sample size for this study was calculated using the Carley et al (2005) nomograms. In order to use the nomograms researchers need to know three of four data sets at the start of the study, namely the number of patients; the CI; the prevalence of the disease in the population; and the level of sensitivity and or specificity. The target population was reached and the narrow confidence intervals confirm that the study had a sufficient sample size to detect effect. With 1313 patient included in the final analysis this is the largest study to date to assess the diagnostic accuracy of the tuning fork test to detect fractures.

4.9.6: Inclusion & exclusion criteria

In order to ensure validity the inclusion and exclusion criteria were defined at the start of the study as recommended by Whiting et al (2003). They were kept to a minimum in order to ensure that the study population was a true reflection of the patients to which the results would be generalised. These were governed by the appropriateness of assessment with the OARs and the age restriction placed on the nurses at the walk-in-centre included in the initial proposal. On withdrawal of the walk-in-centre I could have submitted an ethical amendment to lower the minimum age limit as the other study sites did not have this age restriction for requesting x-rays. However, although this may have increased the population the study findings are applicable to it would have taken time and would have required the development of another age appropriate information sheet.
The study population was a true reflection of the patients in which the test would be used in clinical practice including the severity of the condition, demographics and presence of differential diagnoses. The population was predominantly male, with 54% being aged less than 30 years. The prevalence of fractures nationally in those x-rayed following assessment with the OARs is 15% (Bachmann et al 2003) and the prevalence of fractures in the study population reflected this at 16%. Unfortunately 98% of the study population were of white ethnic origin therefore limiting the ethnic groups the results will be applicable to and resulted in the withdrawal of ‘ethnicity’ as one of the predictor variables in the analysis.

4.9.7: Missing, indeterminate or inconclusive results

Validity of a diagnostic study can be affected by missing data sets (Devillé & Buntinx 2002) and as such the results of the study were first displayed in a 3x3 table that allowed inconclusive, indeterminate and missing data sets to be reported (Sackett & Haynes 2002). Due to the nature of the study no patients were lost to follow-up but there were a number of inconclusive OAR and tuning fork tests. A review of the literature found no other papers reporting inconclusive OARs results so it would appear that this is the first study to report not only inconclusive OAR but also tuning fork test results. All the patients identified as OAR inconclusive went on to receive an x-ray and it was assumed that if the tuning fork test were introduced into clinical practice patients recording
an inconclusive tuning fork test would also receive an x-ray. This ensured consistency in the analysis but further research into what constitutes an inconclusive result is recommended.

Phase two of this study explored the experiences of patients and clinicians involved in the study by collecting qualitative data in the form of focus group discussions. Chapter five describes the methodology and findings of phase two.
Chapter 5 – PHASE TWO: EXPLORATION OF PATIENT AND CLINICIAN EXPERIENCES OF THE TUNING FORK TEST VIA FOCUS GROUP DISCUSSIONS

A secondary outcome measure of this study was to explore patient and clinician views on the tuning fork test and this chapter is divided into two sections. The patient focus groups were undertaken before the clinician focus group and will be discussed in the first section of this chapter which outlines the methodology, the process and the findings from the patient focus group discussions (see section 5.1), whilst the second section of the chapter outlines the methodology, the process and the findings of the clinician focus group discussions (see section 5.2).

5.1: Patient focus group discussions

This section focuses on the aim, objectives, process and findings of the patient focus group discussions.

5.1.1: Aim & objectives

The aim of this section is to answer the following psycho-social question –

‘What is the experience of patients involved in the tuning fork study and what do they view is the potential impact, if any, for patients if the test were introduced into clinical practice?’
The objectives for this phase of the study were therefore to -

1) Explore the initial perceptions of patients when they were approached about the use of a tuning fork to assess an ankle injury

2) Identify descriptors from patients of the sensation felt when the tuning fork was applied to their injured and ‘good’ ankle

3) Identify whether patients believed the tuning fork test to be accurate and whether they thought it would be acceptable to other patients in a similar setting

4) Explore what patients felt were the benefits and disadvantages to patients and the NHS if the tuning fork test was proven to be an accurate method of assessment.

In order to answer the aim and objectives of this phase of the study, a qualitative methodology study was adopted using focus group discussions. Qualitative research is ‘a form of social enquiry that focuses on the way people make sense of their experiences and the world in which they live’ (Holloway & Wheeler 2010, p.3) and involves research that seeks to explore phenomena to understand the experiences, behaviours and feelings from the participants perspective and in their own words. As such it is an appropriate method to meet the aims and objectives of this phase of the study.

Focus group discussions are defined by Sim (1998) as group interviews that focus on a specific topic with the aim of eliciting ideas, thoughts and
perceptions of a lived experience. Through focus groups the researcher can identify the needs and feelings across discussions with a number of participants who, as co-members of the group, are encouraged to join in the discussion. However, while data is produced through social interaction careful moderation is required to ensure all members feel able to disclose new and spontaneous ideas (Kitzinger 2005).

It is difficult in qualitative methodology to estimate the number of participants that are needed to provide sufficient data to describe the patient experiences. However, it was anticipated that no more than six focus group discussions would be required to provide saturation of data or no further new theme development. If however, data saturation was not achieved, the Study Steering Group members agreed that additional focus groups for participants attending study sites C and D could be sought.

As with all research there were ethical considerations for this phase of the study and before any data collection took place full ethical approval was received from the Local NHS Research Ethics Committee as well as the School of Health Science and Social Work Ethics Committee of the University of Portsmouth.
5.1.2: Informed consent

The information sheet given to all patients in phase one of the study included details about the focus group discussions. All patients were invited to consent to being contacted about the focus group discussions as part of the initial consent process into phase one of the study. Patients were reassured that not wanting to take part in a focus group discussion did not affect their participation in phase one of the study and they were informed that their participation was completely voluntary. If a patient declined to take part in the focus group discussion they were not asked for the reason for doing so, however, some patients volunteered that although they were happy to take part in phase one of the study they did not want to come back to the focus group discussions as they did not want to commit to attending due to time and work pressures.

When the patients attended a focus group discussion they were required to give their written consent after reading an information sheet specific to this phase of the study (Patient information sheet and consent form - appendix E). The information sheet reminded the patient about the aim of this phase of the study, informed them what taking part involved, and reminded them that the discussion would be audio recorded. The patients were reassured that their inclusion into the focus group was completely voluntary and that they could withdraw at any point. The moderators were responsible for ensuring that all patients had read the information sheet and signed the consent form prior to the start of the focus group. The facilitator again reminded them that the
conversation would be recorded at the start of the discussion, giving them a further chance to withdraw from the study if they wanted to.

Patients were informed that they would be referred to by a number rather than by name in the transcription to maintain confidentiality and no names would be used in any report about the study but that they would be referred to as ‘Male 1’, ‘Female 1’ etc. In order to maintain confidentiality all consent forms and typed transcriptions were stored in a locked filing cabinet, accessible only by the research team. The recordings of the focus groups were deleted after checking the transcriptions for accuracy as per ethical approval.

5.1.3: Recruitment

The patient focus group discussions ran concurrently with phase one of the study and only patients who gave their consent to being contacted about the focus groups at their initial visit and attending study sites A and B were eligible to take part. The main reason for this was that study sites C and D only began data collection in late June 2010 and the dates for the first four focus groups had already been arranged. It was felt that once data had been analysed from the scheduled focus groups a decision would be taken by the Study Steering Group as to whether any new data would be created by holding further focus groups for patients attending study sites C and D.
The initial study protocol indicated that patients would be contacted about the focus groups within three months of their initial visit by telephone. However it soon became apparent that despite the research nurse telephoning patients at various times throughout the day, this method was restricting the study population being contacted. A decision to submit an ethical amendment to contact patients by letter was made. After receiving approval for this ethical amendment all patients who consented to being contacted about the focus groups were sent an invitation letter to a focus group. A stamped addressed envelope was attached to the invitation letter (Invitation letter to focus group - appendix F) along with a tear-off slip to be returned to the research team. The tear-off slip asked patients to indicate which of the following options they would like to choose

i) To attend one of the listed focus groups (ticking their preferred option)

ii) That they were unable to attend one of the focus groups but consented to being contacted about future dates, or

iii) That they wanted to withdraw their consent to take any further part in the study

The invitation letter also included a statement informing the patient that if they were unable to attend a focus group they could post a comment on the forum of the study website (Tuning fork 2011). Figure 21 shows a screen shot of the study forum webpage.
In addition, a copy of the main patient information sheet (Appendix A) was included to remind patients about the main objectives of the study. Once the study administrator received the completed form a place was reserved for the patient at the focus group on the selected date.

5.1.4: Process

All participants recruited at study sites A and B were eligible to attend a focus group discussion. The first four focus group discussions were held at a local hotel that was centrally located to all invited participants. The focus groups were planned for different days of the week and at different times of the day to
try and ensure maximum availability to patients. When data from the first four focus groups was analysed it revealed that no new data was being generated so one further focus group was arranged to confirm that data saturation had been reached. This was held at another centrally located hotel. Table 24 summarises the date and times of the five focus groups.

<table>
<thead>
<tr>
<th>Date</th>
<th>Day of the week</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>19th October 2009</td>
<td>Monday</td>
<td>11am</td>
</tr>
<tr>
<td>23rd October 2009</td>
<td>Friday</td>
<td>5pm</td>
</tr>
<tr>
<td>17th November 2009</td>
<td>Tuesday</td>
<td>11am</td>
</tr>
<tr>
<td>26th November 2009</td>
<td>Thursday</td>
<td>5pm</td>
</tr>
<tr>
<td>16th August 2010</td>
<td>Monday</td>
<td>8pm</td>
</tr>
</tbody>
</table>

Light refreshments were available throughout the focus groups in the form of hot and cold non-alcoholic beverages and sweet pastries. In addition all attendees were given a £10 book voucher as a small token of appreciation for taking part and the option to claim travel expenses to and from the venue.

5.1.4.1: Facilitation

All the focus groups were led by a facilitator whose role it was to manage time effectively, and to ensure that there was an opportunity for everyone to share their views (Sim 1998). The facilitator was a senior lecturer at a university who
is an experienced qualitative researcher with a nursing background. The facilitator had no prior contact with the participants and had no prior knowledge of ankle assessment, the OARs, or on the use of the tuning fork test to detect fractures.

In addition the research nurse and study administrator attended the focus groups to act as moderators. They were responsible for the layout of the room; meeting patients as they arrived; ensuring informed consent was obtained and to answer any clinical questions that arose during the focus group. In addition, the research nurse and study administrator were required to make notes of the main topics discussed in case the digital recorder malfunctioned.

As a senior nurse who works across study site A and B, I have a keen interest in the diagnosis and management of minor injuries and in my current role as a consultant nurse I am particularly keen to find ways of improving patient care and safety whilst at the same time being mindful of the need to reduce costs. Although I had recruited patients into phase one of the study I felt that as the chief investigator of the study I may have inadvertently biased the discussion if I attended the focus groups. I therefore took the decision not to attend any of the focus group discussions.

The focus group discussions ran concurrently with the tuning fork test before any data analysis had taken place so neither the facilitator nor the moderators
knew whether the tuning fork test was accurate and therefore could not influence the subsequent discussions.

5.1.4.2: Interview Schedule

In order to stimulate discussion around focused topics the facilitator worked with an interview schedule which included a set of open-ended questions (Patient interview schedule - appendix G). Key questions for patients were how they felt when they were first approached about the tuning fork test, what sensation they felt when the tuning fork was applied to their ankle, and to identify whether they thought receiving an x-ray or a reduction in waiting time was the most important aspect of attending an emergency care setting.

5.1.4.3: Data collection

All the focus groups were audio-recorded with a digital Olympus WS 110 digital voice recorder purchased specifically for the study. After each focus group I listened to the recording, not to listen to early themes, but to ensure the recording was clear and audible. Although transcribing the data can help the author become familiar with the data (Holloway & Wheeler 2010) I had no experience of undertaking this task, and due to time pressure it was decided to use an experienced independent medical secretary using a digital transcription machine to transcribe verbatim each focus group discussion. Once the transcriptions had been completed I checked each one for accuracy against the relevant audio recording before deleting the recordings as per ethical approval.
5.1.4.4: Data analysis

Narrative analysis can be defined as the procedure through which the researcher organises the data elements of the transcripts into a coherent and developmental account (Riessman 2008). There are a number of potential approaches to analysing narrative data and thematic analysis was chosen as it focuses on allowing the researchers to interpret and theorise from the whole story and is useful for novice qualitative researchers like myself (Holloway & Wheeler 2010). Thematic analysis focuses on the contents of the patient story and the meaning within it, and is useful if the story is told in a non-sequential way (Holloway & Wheeler 2010). Braun and Clarke (2006) claim that thematic analysis is a foundational method for analysing qualitative data as it forms the basis of other techniques. The level that data analysis takes place may be semantic, where the researcher does not go beyond the surface and do not look beyond the data, or latent, where the researcher looks at the underlying issues, making assumptions and conceptualisations (Braun & Clarke 2006). According to Braun and Clarke (2006) it does not matter whether data analysis is semantic or latent but it must be consistent throughout.

The key to thematic analysis is to transform data into explicit codes and then into themes or relationships (Ofer 2009) and Braun and Clarke (2006) describe a six-step process of data analysis which includes familiarisation, generation of initial codes, search for themes, review themes, define and name themes and produce the report. Table 22 includes a summary of each of these steps. In
order to identify the process of the six phases undertaken for this study each is described in detail after table 25.

Table 25: A summary of the phases of thematic analysis as described by Braun & Clarke (2006)

<table>
<thead>
<tr>
<th>Phase</th>
<th>Descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Familiarizing self with data</td>
</tr>
<tr>
<td>2</td>
<td>Generate initial codes</td>
</tr>
<tr>
<td>3</td>
<td>Search for themes</td>
</tr>
<tr>
<td>4</td>
<td>Review themes</td>
</tr>
<tr>
<td>5</td>
<td>Define and name themes</td>
</tr>
<tr>
<td>6</td>
<td>Produce the report</td>
</tr>
</tbody>
</table>

Phase 1 – Familiarising self with data

I listened to the recording of each focus group before they were transcribed to check for clarity of the recording. The recordings were then transcribed by an experienced medical secretary before being sent back to me. When the transcriptions were returned I checked them for accuracy against the original recording, and to ensure that no patient identifiers had been included in the transcriptions. Patients were referred to as ‘male 1’, ‘male 2’, ‘female 1’, ‘female 2’ etc., and each line of text within the transcription was numbered in order to make it easy to refer to important text when cross-checking themes and discussing the findings. The recordings were then erased.
**Phase 2 – Generate initial codes**

After familiarisation with the data the facilitator and I independently identified interesting words, sentences and whole phrases as selected ‘units of meaning’ from the transcribed dialogue. Although I tried to use the computer software package MaxQDA I found it easier to use pen and paper to do this. I went through each transcript using a highlighter pen, allocating them codes until I had marked the data throughout the whole data set. Text deemed not relevant to the topic in question were deemed as ‘dross’ and were excluded from the analysis. Examples of the ‘dross removed’ are “I watch television in the evenings, Strictly being one of my favourites” (Female 2) and “I holiday on a barge, on a narrow boat, that’s about it, I’m retired” (Male 1).

**Phase 3 – Searching for themes**

The facilitator and I then independently grouped the codes with similar meaning into initial themes. I did this by cutting out the highlighted text from the transcription and grouping text with similar content together.

**Phase 4 – Review themes**

The facilitator and I then met to compare similarities and differences in themes developed independently.

**Phase 5 – Define and name themes**
I then aggregated the initial themes into key themes ensuring that they reflected the participant’s experience. These were then cross-checked for accuracy by the facilitator.

Phase 6 – Produce the report

I wrote the report of the findings and selected the quotes from the transcript that best illustrated the theme described. In the report quotes were written verbatim in italics and enclosed within a text box. A run of full stops was used where text had been removed that was not relevant to the theme being discussed or where speech had been interrupted. The quotes were referenced to which focus group they came from, and what gender the participant was. The quote selected was chosen on the basis that it best illustrated the theme being described.

5.1.5: Findings

The following sub-sections outline the recruitment, patient demographics and findings of the patient focus group discussions.

5.1.5.1: Recruitment

The first 345 patients recruited from study sites A and B were eligible for inclusion in the first four focus groups. Of these 234 (68%) did not consent to being contacted about the focus groups at their initial visit. One patient, who
was willing to take part, had been on holiday from the USA so it was deemed not practical to contact her to attend. This patient was sent the details of the study website and encouraged to post comment to the on-line forum. Figure 22 summarises the response rate for the first four focus groups.

A total of 110 invite letters were sent out. The response rate was high with 89 replies (81%). However, of these

- 35 (39%) withdrew their consent for being contacted about the focus groups
- 24 (27%) could not make the planned dates but were interested in being contacted about further ones if arranged
30 (34%) stated they would like to book a place at one of the focus groups.

However, despite reminder letters being sent out two weeks before the focus groups only seven of the thirty participants who had initially booked a place attended the focus groups, this equates to only 23% of those who had confirmed their attendance.

Ethics approval dictated that patients had to be contacted within three months of their initial attendance and 213 patients met the criteria for being contacted about the fifth focus group. Of these only 94 (44%) had consented to being contacted about the focus group discussions. However, of these 55 (59%) lived outside of the area so it was deemed not practicable to invite them to return to the area to attend a focus group, and two had already been excluded from the study for not meeting the inclusion criteria. 37 invitation letters were subsequently sent out. Eleven replies were received equating to a 30% response rate and another was received ‘returned to sender’ unopened. Of these replies, two withdrew their consent, three could not make the date set and six confirmed they would attend the focus group. Two of the patients initially confirming their attendance rang the research nurse on the day of the focus group apologising for having to cancel their attendance due to unforeseen circumstances. A summary of the responses received for the fifth focus group can be found in figure 23.
Figure 23: Summary of response rate for 5th focus group

Therefore a total of 147 invite letters were sent out and 100 replies received, excluding the reply received unopened and marked ‘return to sender’. This equated to a response rate of 68%. Only a total of 10 participants attended across the five focus groups held.
5.1.5.2: Patient demographics

A total of four males and six females attended the focus groups. The mean age was 59 years (range 15-77), which is higher than the mean age in phase one of the study which is reported as 34 years (range 12-91). Despite all patients having access to the study website only two posted comments to the on-line forum. Table 26 contains the patient characteristics from each of the focus groups and the on-line forum.

<table>
<thead>
<tr>
<th>Focus group</th>
<th>Gender</th>
<th>Study number</th>
<th>Age (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>89</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>127</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>191</td>
<td>64</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>84</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>225</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>50</td>
<td>65</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>106</td>
<td>70</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>675</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>692</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>700</td>
<td>68</td>
</tr>
<tr>
<td>Forum 1</td>
<td>F</td>
<td>Not known</td>
<td>Not known</td>
</tr>
<tr>
<td>Forum 2</td>
<td>F</td>
<td>Not known</td>
<td>Not known</td>
</tr>
</tbody>
</table>

Key: F: female; M: male

Due to the nature of the on-line forum it was not possible to identify which patients posted comments onto the forum, therefore age, study number and x-ray results of these patients are not known, but it was possible to distinguish what gender the patients were and the results of the tuning fork test. Six of the
patients attending the focus groups had been diagnosed with a fracture of which two did not receive any further management. Interestingly when the data collection sheets from phase one of the study were reviewed clinicians and patients did not agree on the result of the tuning fork test in three cases (see table 27).

Table 27: Patient & clinician interpretation of the tuning fork test & the x-ray result

<table>
<thead>
<tr>
<th>Focus group number</th>
<th>Patient Age (years)</th>
<th>Clinician interpretation of Tuning fork test</th>
<th>Patient interpretation of Tuning fork test</th>
<th>X-ray result</th>
<th>Plaster cast</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>62</td>
<td>Negative</td>
<td>Negative</td>
<td>NAD</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>Negative</td>
<td>Positive</td>
<td>Fracture</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>64</td>
<td>Negative</td>
<td>Negative</td>
<td>Fracture</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>37</td>
<td>Positive</td>
<td>Negative</td>
<td>Fracture</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>Positive</td>
<td>Negative</td>
<td>NAD</td>
<td>n/a</td>
</tr>
<tr>
<td>3</td>
<td>65</td>
<td>Negative</td>
<td>Negative</td>
<td>Fracture</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>70</td>
<td>Positive</td>
<td>Positive</td>
<td>Fracture</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>77</td>
<td>Negative</td>
<td>Negative</td>
<td>NAD</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>63</td>
<td>Negative</td>
<td>Negative</td>
<td>NAD</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>68</td>
<td>Positive</td>
<td>Positive</td>
<td>Fracture</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Key: NAD = nothing abnormal detected; n/a = not applicable

Two patients attending the focus groups recalled they were ‘tuning fork negative’ but had been recorded as positive by the clinician, and one patient reported they had a positive tuning fork test but was recorded as negative by the clinician.
5.1.5.2: Emergent themes

A total of 235 units of meaning were identified in the text and these were coded and grouped into 24 initial themes which can be found in table 28.

<table>
<thead>
<tr>
<th>Number</th>
<th>Initial Themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mechanism of injury</td>
</tr>
<tr>
<td>2</td>
<td>Self-care and analgesia</td>
</tr>
<tr>
<td>3</td>
<td>Time from injury to attending ED</td>
</tr>
<tr>
<td>4</td>
<td>Transport to ED</td>
</tr>
<tr>
<td>5</td>
<td>Prior knowledge of attending ED</td>
</tr>
<tr>
<td>6</td>
<td>Initial impressions when approached about the tuning fork test</td>
</tr>
<tr>
<td>7</td>
<td>Thoughts of accompanying relatives</td>
</tr>
<tr>
<td>8</td>
<td>Information received in ED</td>
</tr>
<tr>
<td>9</td>
<td>Sensation to ‘good’ leg</td>
</tr>
<tr>
<td>10</td>
<td>Sensation to injured leg</td>
</tr>
<tr>
<td>11</td>
<td>Need for a comparison</td>
</tr>
<tr>
<td>12</td>
<td>Impression of accuracy of the test</td>
</tr>
<tr>
<td>13</td>
<td>Waiting time in ED</td>
</tr>
<tr>
<td>14</td>
<td>Waiting for x-ray</td>
</tr>
<tr>
<td>15</td>
<td>Concerns about x-ray exposure</td>
</tr>
<tr>
<td>16</td>
<td>Risk of x-rays for children and in pregnancy</td>
</tr>
<tr>
<td>17</td>
<td>Mistrust in x-rays</td>
</tr>
<tr>
<td>18</td>
<td>Reducing waiting times</td>
</tr>
<tr>
<td>19</td>
<td>Reducing NHS costs</td>
</tr>
<tr>
<td>20</td>
<td>Reducing travelling time</td>
</tr>
<tr>
<td>21</td>
<td>Use by GPs</td>
</tr>
<tr>
<td>22</td>
<td>Use by other healthcare professionals</td>
</tr>
<tr>
<td>23</td>
<td>Pain on removing plaster</td>
</tr>
<tr>
<td>24</td>
<td>Pain after injury</td>
</tr>
</tbody>
</table>

Similar codes were then aggregated into the four key themes –
• Key theme 1: Perception of injury and prior knowledge of attending ED,

• Key theme 2: Preconceived ideas about x-rays

• Key theme 3: Perception of the tuning fork test

• Key theme 4: Potential advantages and recommendations for future practice

Table 29 shows how the initial themes were aggregated together to make the four key themes. Each key theme is described below.

<table>
<thead>
<tr>
<th>Number</th>
<th>Key theme</th>
<th>Initial theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Perception of injury and prior knowledge of attending ED</td>
<td>Includes mechanism of injury, self-care and analgesia prior to attending ED, time from injury to attending ED, transport to ED, prior knowledge of attending ED, perceived waiting time in ED, progress of the injury, and pain since removal of POP</td>
</tr>
<tr>
<td>2</td>
<td>Preconceived ideas about x-ray</td>
<td>Includes waiting for x-ray, risks of over exposure and to the young and the pregnant, and mistrust in x-rays</td>
</tr>
<tr>
<td>3</td>
<td>Perception of tuning fork test</td>
<td>Includes information received prior to application of the tuning fork test, initial impression, views of relatives / friends, sensation felt to injured and ‘good’ ankle, importance of comparing both sides, and impression of accuracy.</td>
</tr>
<tr>
<td>4</td>
<td>Potential advantages and recommendations for future practice</td>
<td>Includes reduced waiting times, ED queues, costs, x-ray exposure and travelling time, and use by other healthcare practitioners</td>
</tr>
</tbody>
</table>

5.1.6: Summary of findings

The following section summarises the findings, listing them under the four key themes.
5.1.6.1: Key theme 1 – perception of injury & prior knowledge of attending ED

This theme focuses on the mechanism of injury, self-care prior to attending ED, time from injury to attending ED, transport to ED, prior knowledge of attending ED, perceived waiting time in ED, progression of the injury, and pain since removal of the plaster cast.

In order to help the patients relax each focus group commenced by asking them to explain how they had injured their ankle. Recruitment into the study was restricted to patients who presented following simple twisting injuries and therefore all described mechanisms such as falling off kerbs, slipping off pavements and going down stairs. All the patients had sustained inversion injuries to their ankle. Only one patient attending the focus groups injured their ankle playing sport, whilst another had fallen whilst not wearing her special orthotic shoes used to help with her balance.

I was just walking, I didn’t have these shoes which I wear orthotics in because I thought they were clumpy, I didn’t have Victoria Beckham heels on neither and I just twisted my foot like that and it came up in a balloon straight away and bruised .... I am so used to these orthotics though I think they keep me steady. Cause, one physio told me I was flat-footed, which I didn’t appreciate

Focus group 5: female

Two of the patients went straight to hospital after the injury, one of which was taken direct to hospital from the sports ground by ambulance. The remaining patients stated they waited a minimum of 12 hours and a maximum of 3½ weeks before seeking medical advice, relying on family and friends to transport...
them to the study sites when they did so. Two patients had been sent to the ED for x-ray after visiting their GP, with one of these stating that although she attended the GP surgery she did not even get to see the GP before being redirected to the ED for x-ray.

The patient waiting 3½ weeks from injury to attending ED only sought medical advice as he was going on holiday the following week and he was concerned as his ankle was still swollen. He was surprised to hear that he had not only fractured his ankle but also his foot.

I went up a flight of steps to put a catalogue in a door, stepped backwards, went down the steps, fell down the steps, thought I’d twisting my ankle, it was painful but not really painful, walked around on it for 3½ weeks. Erm,[pause] I was going on holiday on my boat the following week and I thought well the swelling has still not gone down, call in A&E, told them a few lies in there because they won’t see you if it’s not within 2 weeks, so I said I’d only done it 2 weeks ago..........when he showed me the x-rays and I couldn’t believe it, the ankle had just cracked straight across.

Focus group 1: Male

A number of patients had attempted to manage their own injury using a combination of the first-aid measures of ice, cold water, strapping and elevation. Two patients who had applied a support bandage prior to seeking medical advice were surprised when they were told by staff at the ED that strapping was no longer recommended following an ankle injury.
Only two patients had taken analgesia prior to attending the ED and each had only taken one dose on the day of the injury. When their ankle was still painful they decided to attend ED prior to taking further analgesia. All of the patients had prior knowledge of attending ED, either as a patient or whilst escorting relatives, and all stated they expected a long wait of at least 3-4 hours. Most were pleasantly surprised when they waited a lot less than this with the majority being discharged within two hours of arrival.

Patients described the time they waited for the initial assessment as the longest wait they experienced, with only one stating that they waited a long time for the x-ray. Most of the patient’s stated that their injuries had settled by the time they attended the focus group, however two described in detail the increase in pain they experienced after the plaster cast had been removed. One female patient summed this up as...
5.1.6.2: Key theme 2 - Preconceived ideas about x-rays

This theme focuses on patient perceptions about x-rays and includes the waiting time for x-ray, risks of over exposure, risks to the young and the pregnant, and mistrust in x-rays.

All the patients identified they were aware of the risk of repeated exposure to x-rays and in particular of the risks to the young and during pregnancy, and all stressed the importance of trying to reduce the number of x-rays that are taken.

In addition two patients highlighted that they did not trust x-rays as they had both had fractures missed before and it was not until they re-attended with injuries years later that they had been told they must have broken the bones...
previously. One patient in particular had had several missed fractures and was sceptical about the accuracy of x-rays.

One patient had initially been informed that he did not have a fracture and was surprised when two days later he got a telephone call from the hospital to say that a fracture had been missed. Although the management for the injury did not change he was pleased to hear from the research nurse at the focus group that all x-rays at the study sites were checked by a radiologist but he was concerned that the same could not happen with the tuning fork test if its future use was adopted.

5.1.6.3: Key theme 3- Perception of the tuning fork test

This theme focuses on the perception of the tuning fork test and includes information and impression of the tuning fork test prior to its application, the views of accompanying relatives and friends, the sensation felt to injured and
'good’ ankle, and the importance of comparing both sides, together with the patients impression of accuracy.

‘Intrigue’, ‘surprise’, ‘curiosity’, and ‘bizarre’ were just some of the words patients used to describe how they felt when they were initially approached about the study. One female patient stated she took part as she knew it wasn’t going to ‘cause her any pain or inconvenience’, stating that ‘new things have to be tested’.

The patients had mixed views on the theory behind the tuning fork test. Although the majority of the patients did not understand the logic behind the test one patient explained that she understood why the tuning fork would be beneficial in detecting fractures as her daughter was a musician.

I can understand the logic in that, my daughter’s a musician so I’ve noticed that if the note changes obviously there’s something that’s not connected, so that was very logical to me.

Focus group 2: Female

Whereas another patient did not understand the logic of the test but was just happy he was going to be seen.

....I couldn’t understand the logic but I was just pleased to be getting some service [laughs] rather than wait for another 3 weeks for a doctor’s appointment because I originally thought that’s what I might have to do

Focus group 1: male
One female patient, who had returned to teaching yoga only six weeks after her fractured ankle, stated she was happy to take part in the study as she thought the tuning fork was linked to ‘healing with vibrations’ but now realises that the test ‘is to do with diagnostics rather than healing’. Only two patients expressed that their first thoughts when approached about the study were ‘was it going to be painful’, with both stating that when they were shown how the tuning fork would be applied their fears were alleviated.

All the patients who had relatives accompanying them to the ED stated that they had been given full support by them to take part in the study, with the majority stating that their relatives had also read the information leaflet. All the patients expressed the importance of receiving information about what to expect with any procedure or test before it is carried out in order to alleviate fears. In particular one female patient stressed that she liked to gain information from the internet before attending hospital for anything so that she would know what she could expect, but stated that with an injury you ‘don’t have that choice do you’.

I like as much information as possible, I know Google and Wikipedia, they may not be accurate but I do like to make sure.....but when it’s an emergency you don’t have that choice do you?

Focus Group 4 female

With reference to the tuning fork test the majority of patients felt they received adequate information about what to expect and that the information sheet was clear. However, one patient also wanted to be told what sensation she would
feel when the tuning fork was applied but could understand why this information was not made readily available prior to the application of the tuning fork.

The majority of patients described the sensation from the tuning fork as a ‘slight vibration’, a ‘gentle tingling’, with one patient posting a message on the forum stating that ‘the sensation was similar to a TENS machine’. Another claimed ‘it went off singing’, and a ‘tightening around the ankle’ was how one female patient described the sensation, whilst two other participants stated that all they could feel was ‘cold metal against their skin’. Although none of the patients described the sensation as being painful or uncomfortable, one suggested that on a scale of one to ten he would have expected to rate the pain from the tuning fork as a seven or eight, but was surprised when what he experienced he would have rated half the score anticipated.

During the study the tuning fork was applied to both ankles and all the patients attending the focus groups could see the logic of this, even though three could not actually remember having the tuning fork test performed on both ankles,
two of these put it down to the fact that they may have ‘blotted it out’ due to the pain they were in with the injury.

In contrast, one male patient explained that he thought the comparison of both ankles was for the benefit of the practitioner rather than for the patient.

It was anticipated at the start of the study that a ‘positive’ tuning fork test would result in an increase in the intensity of the vibrations in the presence of a fracture when applied to the injured ankle. However, patients reported contrasting experiences, with some stating that the vibrations were stronger on the ‘good’ ankle and some that the vibrations were stronger on the injured ankle. Some patients did not actually feel anything at all on the injured ankle with some also questioning whether swelling would affect the accuracy of the test.
Chapter 5: Phase two: exploration of patient and clinician experiences of the tuning fork test via focus group discussions

5.1.6.4: Key theme 4- Potential advantages & recommendations for future practice

This theme focuses on what patients consider are the potential advantages and benefits for patients and the NHS if the tuning fork test is deemed to be accurate. Suggestions include a reduction in exposure to x-rays, waiting times, travelling times, queues in ED, and costs to the NHS.

When asked about whether they thought the tuning fork test was accurate the majority of patients thought it was, even though it became evident to them throughout the focus group that each of them had experienced slightly different reactions to the tuning fork. They expressed that the tuning fork could not be used on its own but should be used as a guide as to whether to x-ray or not and that patients needed to have confidence in it to trust it.

I didn’t feel a thing, couldn’t feel anything at all. They put it on the ankle and above the ankle and I didn’t feel anything at all at either of those places. When she put it on the other ankle, the good ankle, then I could feel it vibrating.

Focus group 1: female

When he did it to me he put it on my good ankle first to show that there was nothing there and then he put it on my bad ankle so that I could see the difference, or feel the difference and I did notice the difference............. I knew what to expect, I know the good ankle, no reaction whatsoever so I knew if it was the same on that side, I hadn’t broken my ankle but as soon as it started singing, I explain it as singing but you don’t hear anything, it’s just a slight vibration

Focus group 1: male
Chapter 5: Phase two: exploration of patient and clinician experiences of the tuning fork test via focus group discussions

Only one patient stated that he would still like to have an x-ray in the presence of a negative tuning fork test. This patient was a 15 year male.

Patients did however question whether swelling would affect the result and whether the effect would be the same on children due to the fact that their bones are less dense.

A reduction in time, treatment costs and exposure to x-rays were the three main benefits patients identified if the tuning fork were introduced into practice.

Travelling to one of the study sites is notoriously difficult at certain times of the day and reducing time was not only discussed in terms of reduced waiting times in the ED and for x-ray but also of the inconvenience of travelling to this site.
Chapter 5: Phase two: exploration of patient and clinician experiences of the tuning fork test via focus group discussions

One male patient felt very strongly that the tuning fork test should be introduced into GP practices to avoid going to the ED in the first place and also questioned whether the tuning fork test could be used to assess healing at the time the plaster cast is removed.

Although other patients agreed on the need for GPs to be trained in the use of the tuning fork test they discussed the logistics of actually getting a GP appointment. Issues raised included not being able to get an appointment at all, getting an appointment but then not being seen by the GP, and also that it was more convenient to go to ED than wait to see a GP.
Since taking part in the study patients had recommended the use of the tuning fork to their GPs and one to her ‘podiatrists’. In addition, another had mentioned the tuning fork study to her son, who was a ‘vet’.

5.1.7: Discussion of findings

This is the first study to explore patient perceptions on the use of the tuning fork test to detect fractures. The aim of the focus group discussions was to explore the experience of patients when they received the tuning fork test to assess their ankle injury but numbers attending were low and did not truly reflect the demographics of the participants in phase one of the study. 63% (353/558) of patients eligible to attend did not consent to being contacted about the focus groups, limiting the number of patients available. Although the mean age of patients attending the focus groups was 59 years (range 15-77) eight out of the ten patients attending the focus groups were aged between 62 and 77 years. This is in contrast to the main study population in phase one, where only 8% were over 60 years of age and means that the younger population which made up the majority of the main study population were unrepresented at the focus group discussions. Furthermore, all patients attending the focus groups were of white ethnic origin and as such the small number of non-white ethnic groups recruited into phase one of the study were unrepresented.

Key theme two focused on the patient’s preconceived ideas about x-ray and although x-rays are deemed to be the gold-standard test to diagnose fractures
in the acute setting comments made by two of the participants suggest that not all patients trust them. Both these patients recalled experiences of having fractures missed by x-rays in the past and patients were surprised to hear that all x-rays were checked by a radiographer as part of routine clinical practice. Patients raised concerns that this checking procedure could not happen with the tuning fork test if it were introduced into clinical practice in the future.

All the participants were aware of the risks of repeat exposure to x-rays, the risk of x-rays to the young and the risks during pregnancy, and along with reducing waiting time and costs, rated a reduction in x-rays as one of the main benefits of the tuning fork test. None of the patients claimed they expected an x-ray on arriving to the ED, supporting the work of Anis et al (1995) who found that patients were equally satisfied with the care they received whether they received an x-ray or not, and concluded that patients do not present to ED with the preconception that they want an x-ray.

Patient descriptors under key theme three suggest that the tuning fork is acceptable to the majority of the patients, even though the majority were initially intrigued and surprised on hearing about the tuning fork test. Provided they received information the majority of the patients gave the impression they would willingly accept the tuning fork as a method to assess fractures of the ankle following injury. However, all patients in this study knew they would receive an x-ray regardless of the tuning fork test and as such this may have biased their opinion. When questioned only one patient, a 15 year old male, said he would
still like to receive an x-ray in the event of a negative tuning fork test just to confirm there was no fracture. Although this is only one patient it can be argued that this is a voice of a younger study participant and therefore reflects the results of phase one of the study where age was seen as a predictor variable of a accurate tuning fork test.

The clinicians had all received training on how to apply the tuning fork test and had been told to look for signs of discomfort from the patients to indicate a positive test. This is in keeping with the previous studies on the use of the tuning fork to detect fractures by Lesho (1997), Van den Berg (2003), Dissmann & Han (2006), and Wilder et al (2009), who all claim the tuning fork test causes a degree of discomfort in the presence of a fracture. However, none of the patients described the tuning fork test as being painful and some actually described a reduction in the sensation when the tuning fork was applied to their injured ankle compared to their ‘good’ ankle. This has implications for introducing the tuning fork test into clinical practice. Instead of a sign of discomfort indicating a positive tuning fork test it may be that a difference between the injured and ‘good’ ankle may need to be taken into account. The comments by the majority of patients in key theme three suggested they could see the logic of comparing both sides.

Although the patients all stated they thought the tuning fork was accurate three of them recalled different results to that recorded by the clinician. One of these was the patient who presented 3½ weeks post injury. He was recorded as
falsely negative by the clinician, while he recalled he had a positive tuning fork test. In addition, both participants at the second focus group thought they were tuning fork negative but had been recorded by the assessing clinician as positive. One of these was the 15 year old male who had stated he would still like to receive an x-ray in the event of a negative tuning fork test. He had been recorded as positive by the clinician but rated himself as negative. As the results of the tuning fork test were not known at the time the focus groups were held it was not explored further as to whether this would have made a difference to his impression on the need to x-ray in the event of a negative tuning fork test.

Apart from personal acceptability of the tuning fork test four patients under key theme four shared their views on the use of the tuning fork in everyday clinical practice, suggesting whether it could be used in children, on different fractures types, and in keeping with Dissmann and Han (2006) whether swelling would have an effect on the result of the tuning fork test. Dissmann and Han (2006) did not include patients with swelling in their study and it would appear that this study is the first to do so. Analysis of swelling as an independent variable to an accurate tuning fork test in phase one of this study found that the tuning fork test was more accurate in the presence of significant swelling than no swelling (see section 4.6.3), and this again has implications for the introduction of the tuning fork into clinical practice. It is not known why this is the case and further research in this area is needed.
Under key theme four patients rated a reduction in x-rays, a reduction in waiting times and costs to the NHS as the main benefits for introducing the tuning fork test into clinical practice. However, despite the concerns they had about the logistics of actually getting an appointment to see a GP within a reasonable length of time some patients made the recommendation that the tuning fork test should be extended into GP surgeries. They suggested that by the time they had travelled to the ED it was probably too late for the tuning fork test to be used, quoting time spent travelling and the fact that x-ray facilities were readily available in the ED as reasons for not implementing the tuning fork test into ED.

5.1.8: Conclusion

This study appears to show that with careful explanation the tuning fork test is acceptable to the majority of the patients attending. However, the younger population are not represented. In order to be introduced into clinical practice the tuning fork test also needs to be seen as reliable and acceptable to the clinicians that will be using it and the following section describes the process of exploring the experiences of the clinicians involved in phase one of this study with the aim of identifying whether the tuning fork test is acceptable to clinicians.

5.2: Clinician focus group discussions

As well as identifying the experience of patients involved in the study it was also important to identify what clinicians involved in the study felt about the tuning
fork test, as any test introduced into clinical practice has to be acceptable to both patients and clinicians. Furthermore, no other study on the use of the tuning fork to detect fractures had explored clinician experience of using it in clinical practice.

5.2.1: Aims & objectives

The aim of focus group discussions with clinicians was to answer the following psycho-social question –

“What is the experience of clinicians involved in using the tuning fork test to diagnose ankle fractures following simple twisting mechanisms, and do they view it as an acceptable method of assessment?”

The objectives for this phase of the study are therefore to -

1. Explore what clinicians thought when they were first approached about using the tuning fork test.

2. Identify whether clinicians felt the tuning fork test was an accurate method to examine an ankle.

3. Identify whether clinicians felt the tuning fork test could be used by other healthcare professionals.

4. Explore what clinicians felt were the potential benefits and disadvantages to patients and the NHS if the tuning fork test was introduced into clinical practice.
In order to answer the aims and objectives above, a qualitative methodology study using focus group discussions was adopted, as per the methodology for the patient focus groups. In order to hold the clinician focus groups a substantive amendment to the ethical approval was obtained from the Local NHS Research Ethics Committee before any data was collected as I felt it was not made clear in the initial protocol that clinicians would be asked to attend the focus group discussions. It was anticipated that two clinician focus groups would provide rich data to describe the experience of involvement in the study, but that after data analysis further focus groups would be held if the data was insufficient.

5.2.2: Informed consent

All clinicians involved in recruiting patients into phase one of the study from all the study sites were eligible to attend the focus group discussions. Clinicians were given an information sheet prior to the start of the focus group which explained the objectives of this phase of the study, that they would remain anonymous in any written report about the study, and that the discussion would be audio recorded. They were then asked to sign a consent form (Clinician information sheet and consent form – appendix I). The facilitator reminded the clinicians that they would not be referred to by name in any transcription or report about the study but would be referred to by gender and number (‘Male: ENP 1’, ‘Female: ENP 2’ etc.), that the discussion was going to be audio recorded and that they could withdraw at any time before the start of the focus
group. As with the patient focus groups all consent forms and transcripts were kept in a lockable filing cabinet which was only accessible by the research team, and the audio recordings were deleted after the transcriptions had been checked for accuracy.

5.2.3: Recruitment

The clinician focus groups were arranged to take place towards the end of the data collection phase for phase one of the study and after the patient focus group discussions but before any analysis had taken place. The clinicians were invited to attend the focus groups by NHS trust email and a poster displayed at each of the study sites. Once data had been analysed from the initial scheduled clinician focus groups it was agreed by the Study Steering Group Committee that a decision would be taken as to whether any new data would be created by holding further clinician focus groups.

5.2.4: Process

The clinician focus groups were held in a meeting room of the emergency department at study site A and took place on the 9th and 22nd March 2011 at 6.30pm and 12.30pm respectively. Light refreshments, including hot and cold non-alcoholic drinks and buffet food, were available throughout the focus groups, and all attendees were given, as a small token of appreciation for
taking part, a £10 book voucher and the option to claim travel expenses. Overtime was also paid if the clinician agreed to attend the focus group on an allocated day off.

5.2.4.1: Facilitation

The same facilitator used for the patient focus groups, was used for the clinician focus groups, and the research nurse and study administrator again acted as moderators. As I work closely with the clinicians involved in the study I did not attend the focus groups as I felt my presence could have influenced the discussion, inhibiting my colleagues from sharing their views and subsequently influencing the direction of the findings. As with the patient focus groups the clinician focus groups ran concurrently with phase one of the study, and everybody, including the facilitator and the moderators did not know whether the tuning fork test was accurate or not and therefore could not influence the subsequent discussions.

5.2.4.2: Interview schedule

In order to stimulate discussion around focused topics the facilitator again worked with an interview schedule which included a set of open-ended questions (Clinician interview schedule - appendix I). Key questions focused on asking clinician’s to share their experiences of involvement in a clinical trial, their feelings about the tuning fork test, whether they felt the test was acceptable to patients as a diagnostic test, and what they felt the benefits or barriers to implementation of the test into clinical practice would be.
5.2.4.3: Data collection & analysis

Data collection and analysis was undertaken in a similar way as for the patient focus group discussions (See section 5.1.4.3: Data Collection and section 5.1.4.4: Data Analysis).

5.2.5: Findings

This sub-section outlines the findings of the clinician focus groups and includes recruitment, and emergent themes.

5.2.5.1: Recruitment

113 clinicians (60 nurses and 53 doctors) were involved in recruiting patients into phase one of this study. The majority of the doctors had been junior doctors on rotation to the departments at the time of their involvement and had since left the study sites. It was therefore difficult to track down the junior doctors to invite them to the focus groups. Invitation to the clinician focus groups was therefore restricted to the nurses and senior doctors involved in recruiting patients into phase one of the study at each of the study sites.

5.2.5.2: Demographics of clinicians

A total of ten ENPs attended the focus groups, all were employed to work on rotation between study sites A and B. Eight were female and two were male.
and all were experienced ENPs, having been employed in the role of ENP for between five and ten years. No doctors attended the focus groups.

5.2.5.3: Emergent themes

A total of 224 units of meaning were identified in the text and these were coded and grouped into 18 initial themes. Table 30 shows the key themes identified.

<table>
<thead>
<tr>
<th>Number</th>
<th>Initial Themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Time in role as ENP</td>
</tr>
<tr>
<td>2</td>
<td>Use and impression of Ottawa ankle rules</td>
</tr>
<tr>
<td>3</td>
<td>Clinical examination by ENPs</td>
</tr>
<tr>
<td>4</td>
<td>Clinical examination by ED doctors</td>
</tr>
<tr>
<td>5</td>
<td>Clinical examination by GPs</td>
</tr>
<tr>
<td>6</td>
<td>Patient expectations</td>
</tr>
<tr>
<td>7</td>
<td>Initial impression of tuning fork test</td>
</tr>
<tr>
<td>8</td>
<td>ENP perception of what patients thought of tuning fork test</td>
</tr>
<tr>
<td>9</td>
<td>Suggested adaptations needed to tuning fork test</td>
</tr>
<tr>
<td>10</td>
<td>Accuracy of tuning fork test</td>
</tr>
<tr>
<td>11</td>
<td>Analgesia and pain</td>
</tr>
<tr>
<td>12</td>
<td>Involvement in research project</td>
</tr>
<tr>
<td>13</td>
<td>Training received</td>
</tr>
<tr>
<td>14</td>
<td>Waiting times</td>
</tr>
<tr>
<td>15</td>
<td>Compliance of patients with management of injury</td>
</tr>
<tr>
<td>16</td>
<td>ENP perceived risks of x-ray exposure</td>
</tr>
<tr>
<td>17</td>
<td>Patients lack of knowledge about dangers of x-ray</td>
</tr>
<tr>
<td>18</td>
<td>Future use of the tuning fork test</td>
</tr>
</tbody>
</table>

These 18 initial themes were then aggregated to make five key themes –
• Key theme one: Experience as an ENP and knowledge of the Ottawa ankle rules

• Key theme two: Comparison of clinical assessment by ENPs and doctors

• Key theme three: Perceptions on patient expectation

• Key theme four: Experience of using the tuning fork

• Key theme five: Recommendations for future use of the tuning fork test.

How the initial themes have been grouped together and aggregated into the five key themes can be found in table 31.

<table>
<thead>
<tr>
<th>Number</th>
<th>Key theme</th>
<th>Initial themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Experience as an ENP and knowledge of the Ottawa ankle rules</td>
<td>Includes time in role as ENP and experience of using the Ottawa ankle rules</td>
</tr>
<tr>
<td>2</td>
<td>Comparison of clinical assessment by ENPs and doctors</td>
<td>Includes ENP thoughts on the differences in the clinical examination performed by themselves, ED doctors and GPs.</td>
</tr>
<tr>
<td>3</td>
<td>ENP perceptions on patient expectations</td>
<td>Includes patient expectations of what to expect in ED, pain and use of analgesia, waiting times, compliance by patients with management of injury and the perception that patients are unaware of the risks of x-ray exposure.</td>
</tr>
<tr>
<td>4</td>
<td>Experiences of using the tuning fork</td>
<td>Includes how ENPs felt about being involved in the research project, the training received, their initial impression of the tuning fork test and their perception of what patients thought of the test, and how accurate they felt the tuning fork test was.</td>
</tr>
<tr>
<td>5</td>
<td>Recommendations for future use of the tuning fork test</td>
<td>Include the recommendation for future use of the tuning fork test and the adaptations they feel are needed to ensure the tuning fork test is acceptable as a clinical assessment tool.</td>
</tr>
</tbody>
</table>
5.2.6: Summary of findings

The findings from the clinician focus groups are outlined below under the key themes.

5.2.6.1: Key theme 1 - Experience as an ENP & knowledge of the OARs

This theme includes the time the clinicians had worked in the role of ENP and their experience of using the OARs. Although one nurse practitioner had only commenced work at the study site the previous week all were experienced nurse practitioners with between five and ten years of experience in the role. All the ENPs had experience of assessing patients presenting with ankle injuries and suggested that they could predict when the patients were going to come in.

All the ENPs knew about the OARs and were able to recite them from memory, but there were differences of opinion in their usefulness in the clinical setting. When asked about whether they used the OARs the majority of the ENPs replied ‘yes’ and ‘absolutely’, stating that the rules were easy to use.

You can guarantee when the ankle injuries are going to come in. Sunday mornings after football and Monday and Wednesday evenings after 5-a-side, you can actually put times and days on it. All the 5-a-siders come in at 10 o’clock at night

Female (ENP 4)
Always use them, you have to don’t you, it’s one of those things, it’s one of the criteria isn’t it?

Female (ENP 9)

I think that Ottawa rules are so good that you don’t need to deviate from them, you use your clinical judgement but you will find that your clinical judgement will actually would match the Ottawa rules therefore because they are, they are, they are some of the few rules that you actually think, they actually really work

Female (ENP 7)

However, in contrast two stated that although the OARs were useful as a guide when starting out as an ENP, they felt if you followed them exactly as you gained experience you would actually request more x-rays. One of these ENPs also claimed that he felt patients were also learning the rules.

I think personally from my experience if you follow them [Ottawa ankle rules] to the letter you know I think they make you x-ray more, from personal experience because I think you gain, I know it’s not a word people like, you gain intuition over a period of time with experience and you tend to know rather than just the pushing, the touching and the feeling bit, rules are great as a guide for a beginner but I think as experience comes you don’t necessarily need to use them..........Patients are also learning the rules I think and this is awful to say, but I think they know when to say ‘Ow’.

Male (ENP 2)

5.2.6.2: key theme 2- Comparison of clinical assessment by ENPs & doctors

This theme includes what the ENPs felt about their own assessment skills and how they compared their own skills to those of doctors working within the emergency department and GPs. All the ENPs were proud of their assessment
skills claiming that, in their view, they were more systematic in their assessment technique than the doctors.

The ENPs put their superior examination skills down to the continual experience gained from consistently managing minor injuries over time, and that they have to clinically justify every x-ray request they requested.

As well as having the perception that their clinical assessments were far more robust than that of the doctors the majority of ENPs agreed, in their opinion that the majority of doctors tended to x-ray everything.

All the ENPs recalled examples where they had assessed patients, informed them that they do not need an x-ray only to find that the doctors had
reassessed the same patient the next day, sent them to x-ray and then diagnosed a sprain.

I have an example that saw a patient with a sprain and it was an absolute, categorical sprain, no bony tenderness whatsoever, sent them away and spent time explaining this is why it hurts, it’s the ligaments that hurt, see I’m pressing here on the bone and it’s not painful and I came in the following day and I saw this woman going down to x-ray, one of our consultants had seen her and I said “so why are you x-raying her?”

Female (ENP 4)

One ENP suggested doctors had a tendency to x-ray more due to the fact that doctors were fearful of litigation and complaints.

It’s also the threat of complaint, ’I didn’t get the treatment I think I need and so’........I mean there’s sort of avoidance from the medical side.....they [patients] will write letters of complaint which then cost money to answer

Male (ENP 2)

It was not only the assessment by ED doctors that the ENPs felt were inferior to their own but also that of GPs. The ENPs had very strong views about the ability of GPs to assess ankle injuries, claiming that the majority have lost the skill to assess minor injury.

I think their [GPs] lack of clinical examination, some of them, it’s very poor....... If they knew how to examine patients for injuries........ they’re not used to examining patients for injury anymore, they have lost that skill

Female (ENP 4)

that’s half our battle because the GP will send them up and say to them, you need an x-ray without properly examining the patient’s ankle and of course they come to us, we properly examine the patient’s ankle and tell them, actually you don’t need an x-ray and they[patient] argue with you and say ’my GP sent me up for an x-ray’.

Female (ENP 3)
In addition, the ENPs were in agreement in stating that it was quicker for a GP to send a patient direct to an ED rather than assess the patient themselves in the surgery, stating that if the patient was assessed as requiring an x-ray the GP would have to send the patient to the ED anyway. One ENP summed this up as...

```
if someone phones up and says I’ve gone over on my ankle, can I have an appointment, they just say, go to the A&E department because that’s our work, it’s not theirs and then if that person needs an x-ray, we can get it done and their treatment can get moved along whereas otherwise they have to go sit and wait in the surgery for the GP to see them, he’ll then say, yes, I think you need an x-ray, he can’t write them up for an x-ray, they have to then come and be seen all over again, so they can’t GPs can’t just send them straight for an x-ray, they have to come to us, we would then sit them in our waiting room, they could be waiting for another 2 hours to be seen all over again....... I think that’s why they just re-direct to us......It’s not that GPs don’t want to see these patients, it’s just that from the patient’s point of view, it’s much quicker to come down and see us
```

Female (ENP 7)

**5.2.6.3: Key theme 3 – ENP perception on patient expectations**

This theme includes what ENPs perceived were patient expectations on arrival to the ED, waiting times, compliance by patients with management of the injury and the perception that patients are unaware of the risks of x-ray exposure.

In contrast to the findings of the patient focus groups but in keeping with the work of Bessen et al (2009) the majority of the ENPs held the view that patients expect and often demand an x-ray when they present with an ankle injury.
Chapter 5: Phase two: exploration of patient and clinician experiences of the tuning fork test via focus group discussions

All the ENPs recalled examples of patients presenting to more than one care setting in order to ‘get the x-ray they want’, claiming that peer pressure may be to blame.

Despite these perceptions the ENPs were convinced that with the correct information patients would have fewer expectations and be more compliant with treatment regimes.
One ENP stressed that the information on ankle injuries should be directed by the wider NHS and not just locally led, and even recommended charging the patients if they demand an x-ray when not clinically indicated.

> I think we should charge them all £10 and they get a reward as a deduction for everything done right [laughs], eventually they’d do it right.......we should charge them, if they demand an x-ray we should charge them.......if we’re wrong give them their money back ........I think if the NHS or MIU, PCTs, GPs were very much more robust saying this is what you will get, and this is what you will not get so do not expect it, and that was broadcast nationally then maybe attitudes would change

Male (ENP 2)

Another ENP, who had just left a walk-in-centre to work at study site A, suggested that x-ray requests could be used in the same way as ‘delayed prescriptions’ as used by GPs for antibiotics. In other words if the patient demands an x-ray, but is deemed not to require one, they are given an x-ray request form to use if the injury is not improving in a few days time. The x-ray would not be taken until the date entered on the form and this particular ENP questioned whether patients would actually bother going for the x-ray if they were put into this situation.

In contrast to the findings from the patient focus groups ENPs had the perception that patients demanded x-rays as they do not understand the dangers of x-ray exposure.
Chapter 5: Phase two: exploration of patient and clinician experiences of the tuning fork test via focus group discussions

5.2.6.4: Key Theme 4 – Experiences of using the tuning fork

This theme includes how ENPs felt about being involved in the research project, what they thought of the training they received specific to the study, their initial impression of the tuning fork test, their perception of what patients thought of the test, use of analgesia, and how accurate they felt the tuning fork test was.

All but one of the ENPs had been involved in recruiting patients into at least one other clinical trial prior to their involvement in this study. When the ENPs first heard about the tuning fork test they thought it was ‘interesting’, ‘curious’, and ‘exciting’, with one ENP claiming that their first thought was about the ‘extra paperwork’ involved.

I just say that you know, you don’t want to be exposed to radiation unnecessarily, but they still say ‘I do’ [laughs]

Female (ENP 9)

they [patients] don’t realise that x-rays are potentially unsafe if you over-expose

Male (ENP 2)

amongst many, oh my god, extra paperwork if I’m [laughs] going to be honest

Female (ENP 5)
Chapter 5: Phase two: exploration of patient and clinician experiences of the tuning fork test via focus group discussions

All the ENPs received training on the tuning fork test, but stated that this training varied. Some of the ENPs had attended group training sessions, whilst others had one-to-ones with me. One ENP claimed she did not receive any formal training, but was shown by an ENP colleague. Despite the different training sessions all the ENPs claimed that the tuning fork test was simple, straightforward and easy to learn and apply.

I think initially people thought it was more complex than what it is...........It was watch one, do one, teach one really, it was that simple

Male (ENP 2)

Only one ENP had sought additional training as she had not fully understood the process and she came back to seek additional help.

she showed me but I had to go and ask her 2 or 3 times and after that . . . . I don’t know, there were so many things to remember and I thought when sometimes I know you can read about things but when shows you a few times I find it easier, so it’s probably just me, cos everyone else was alright I did have to go and speak to her 2 or 3 times and make sure I’d got it right

Female (ENP 9)

Despite stating that the test was easy to learn a couple of the ENPs explained they had reservations that they were using the tuning fork properly. Their main concern was whether they had made the tuning fork vibrate enough and how much pressure they should actually be applying.
Most of the ENPs stated that the majority of patients were interested in the study, were willing and happy to take part, and accepted the tuning fork test as part of the assessment process once the study had been discussed with them.

Only one ENP suggested that, in her view, the patients thought the test was ‘bizarre’ and had lots of patients refuse to take part. Whilst a few of the ENPs revealed they needed to reassure patients that the test did not involve hitting them with the tuning fork and that their treatment was not going to change, that they would still get the x-ray they required.
Chapter 5: Phase two: exploration of patient and clinician experiences of the tuning fork test via focus group discussions

The ENPs described the tuning fork study as being ‘less faffy’ than other clinical trials they had been involved in, and stated that although the tuning fork study was easier to do than other clinical trials it still impacted on the work of the department. They stated that the process of going through the paperwork and getting the consent form signed took longer than the actual test itself, and that time and workload pressures and staffing levels were the main barriers to recruitment.

In contrast to the patients, who had all stated they thought the tuning fork was accurate, all the ENPs expressed that, in their opinion, it was not, suggesting that the tuning fork would miss a lot of fractures if introduced into clinical practice. The ENPs claimed that the patient responses were very ‘random’, and

They think it’s bizarre and you know, even when you show them the tuning fork and then until you have had time to fully explain, they think you’re going to hit them with the tuning fork

Female (ENP 6)

I have had patients think we were going to attack them with the tuning fork......I think you have to almost guarantee them and really reassure them that their treatment isn’t going to change and really reassure them that it’s not going to change anything, it’s in addition to but you do see that, again the intuitive feeling that you get, that they think I’m going to change something and they’re not going to get what they want

Male (ENP 2)

when you’re under so much pressure to get them through, if they [patients] haven’t been given the leaflet in reception and I will hold my hand up to say, to be honest, I just didn’t have time to say ‘do you want to take part in the tuning fork’, that adds another 5 minutes onto the whole assessment process and I just didn’t have time sometimes

Female (ENP 4)

In contrast to the patients, who had all stated they thought the tuning fork was accurate, all the ENPs expressed that, in their opinion, it was not, suggesting that the tuning fork would miss a lot of fractures if introduced into clinical practice. The ENPs claimed that the patient responses were very ‘random’, and
‘a real mixed bag’, with patients diagnosed with and without fractures showing no sign of discomfort when the tuning fork was applied.

```
In my experience I think I can honestly say with every patient I have done it on, none of them whether they had a fracture of not have complained of pain or discomfort, none of them........with the vibration, whether they have got a fracture or not, none of them have displayed or expressed discomfort from the tuning fork.
```

Female (ENP 3)

The Ottawa ankle rules were designed to assist the examination of the ankle and determine the need for x-ray but were not designed to replace clinical judgement in the presence of obvious deformity (Stiell et al 1992) but one ENP commented that she used the rules on a patient who had ‘clearly fractured their ankle’.

```
......and I had patients that had clearly fractured their ankle and when they put the tuning fork on they haven’t displayed any discomfort, haven’t pulled their foot away so I find it quite interesting
```

Female (ENP 3)

She was not the only one however as four patients had been removed from the analysis in phase one of the study due to having obvious deformity requiring manipulation in the ED (see section 4.7.1 Recruitment) and in response to this another ENP questioned the appropriateness of using the tuning fork test on patients with clinically fractured ankle anyway.
Chapter 5: Phase two: exploration of patient and clinician experiences of the tuning fork test via focus group discussions

Even though another ENP agreed with the random results she explained that a positive test usually meant that the x-ray was abnormal in some way, even if this was not always with an acute fracture.

One ENP also questioned whether analgesia had an effect on the tuning fork result, identifying that recording the analgesia a patient had been given had not part of the data collection process of the study.

This comment led to a discussion about patients and analgesia in general with the ENPs suggesting that patients do not like taking analgesia when they have an injury.
5.2.6.5: Key Theme 5 – Recommendations for future use of the tuning fork test

This final theme includes the ENP’s recommendations for future use of the tuning fork test and the adaptations they felt would be needed to ensure the tuning fork test is acceptable as a clinical assessment tool.

Despite the reservations about the accuracy of the tuning fork test the majority of the ENPs stated that, if the study showed it to be reliable, they would be happy to use it as part of their clinical assessment, but they all stressed that it could not be used on its own and would have to be used in conjunction with the Ottawa ankle rules.
The ENPs discussed different groups of patients where assessment can sometimes be difficult and where the tuning fork may be useful as part of their assessment, this included children, and patients with learning disabilities. They also discussed adaptations that could be made to make the test more reliable, such as a light to know that the tuning fork was vibrating correctly. Some of the ENPs had expressed concern that they were unsure whether the tuning fork was vibrating properly and one male ENP suggested a battery-operated tuning fork would produce standard vibrations and could therefore be more sensitive and accurate. He also suggested that a tuning fork with a pointed end rather than a flat end may result in better bone contact than the flat ended tuning fork used for the study.
As with the patient focus groups one ENP raised the issue as to whether the tuning fork test could be used in the pre-hospital setting, for example in GP surgeries, first aid centres and by paramedics. Although the majority of the other ENPs agreed it would be beneficial to patients in rural areas they stated it would not be practical in an inner city setting.

The main reason for this was the concern that patients can not currently obtain appointments with their GPs easily, particularly at weekends and out-of-hours, and that it is easier and quicker for the patients to attend ED.

In addition, one ENP suggested it may be more beneficial to give patients their own tuning fork through ‘letter boxes’ and put them in ‘first aid boxes’ than allow the GPs to have them.
5.3: Discussion of findings

A total of 60 ENPs were involved in recruiting patients and performing the tuning fork test in phase one of the study but only 10 attended the focus group, two male and eight females. All had been working in the role of ENP for five or more years at study sites A and B, and all had received training on the OARs during their ENP training. In keeping with the work by Brehaut et al (2005), all the ENPs had a working knowledge of the OARs prior to the start of the study; with the majority using them regularly in their everyday practice.

Under key theme two ENPs acknowledged that they were the ones who assessed minor injuries on a daily basis and therefore had better skills to assess minor injury when compared to doctors. Reflecting the work of Bachmann et al (2003) the ENPs suggested that fear of litigation was the main reason the doctors had a tendency to x-ray more. This was in contrast to their own practice where they felt obliged, as nurses, to undertake a thorough assessment in order to be able to justify every x-ray they request.

I think if you were going to do that, it would be a better idea just to put tuning forks in letter boxes and tell them to put them in first aid boxes, it would be as effective, I mean, that would be ultimate way of stopping them, if it worked

Male (ENP 2)
Under key theme three the ENPs suggested that doctors have lost the skills to assess minor injuries, leading to an increase in expectation that they will receive an x-ray on their arrival to the ED. This is in keeping with Bessen et al (2009) who concluded that ENPs hold the preconception that patients expect an x-ray when they attend with a minor injury. In addition, Pigman, Klug, Sanford and Jolly (1994) concluded that patients are dissatisfied if they do not receive an x-ray for their injury and all the ENPs could recall examples of patients returning to the ED the following day if they did not get an x-ray on their first attendance. However, this is in contrast to the findings of Anis et al (1995) and Wilson et al (2002) where patients made it quite clear that they do not go to ED with the perception to expect an x-ray. ENPs suggested this perception exists as patients are unaware of the dangers of radiation but this is in contrast to the findings from the patient focus groups where the patients made it clear that they do understand the risks and in fact were adamant that reducing exposure to x-rays was a must for the NHS, quoting this as one of the main benefits of the study (see section 5.1.6.2: Key theme 2 – Preconceived ideas about x-rays).

The patients had all stated that the tuning fork test was not painful and descriptors under key theme four identified that ENPs were in agreement. They stated that in their experience no patient had shown signs of pain, but in contrast to the patients, all the ENPs stated they did not think the test was accurate. The ENPs claimed that the tuning fork test results had been a ‘real mixed bag’ claiming that patients with and without fractures had recorded
positive and negative tuning fork tests. Some of the ENPs explained they had doubts they were using the tuning fork correctly but despite the opportunity to access additional training only one attended more than one training session. The ENPs claimed that the tuning fork could not be used in isolation, would always need to be used in conjunction with the OARs, and in agreement with the patients, stated it appeared acceptable to patients, but only after a good explanation of the test had been given.

Patients and ENPs recommended that the tuning fork test could be useful in situations where clinical assessment is difficult i.e. the young and in those with learning disabilities. However, in contrast to the patients the ENPs could not see the logic in the tuning fork test being made available in GP surgeries due to the GPs lack of ability to assess minor injuries, and time constraints for appointments. Both ENPs and patients were in agreement that there were difficulties in the logistics of actually getting an appointment with a GP within a reasonable time frame, with both groups stating it was quicker and easier to attend ED than the GP.

**5.4: Strengths & weaknesses of this phase of the study**

This phase of the study used a qualitative approach in the form of focus group discussions to explore patient acceptability and clinician experiences of using the tuning fork test to diagnose ankle fractures. Evidence on the use of the tuning fork to detect fractures to the ankle is limited to one other small scale
study (Dissmann & Han 2006) and exploring the experiences of patients and clinicians on its use has not previously been undertaken.

There “is no single way to separate ‘good’ from ‘bad’ qualitative research” (Pope & Mays 2006, p. 87) but there are ways in which validity can be improved. Greenhalgh and Taylor (1997) state that the strength of qualitative research lies in validity or ‘closeness to the truth’ (p. 740), and is judged by transferability of the findings, quality control in the form of an audit trail, and reflexivity. Each of these in relation to this study will be discussed below.

5.4.1: Transferability

It is recognised that findings of qualitative research may only be transferable to the small sample investigated and only then if reliability has been achieved by clearly documenting the study process. Despite efforts to recruit representative focus groups, a large percentage of patients recruited into phase one of the study declined to be contacted about the focus group discussions. This limited the proportion of the main study population available to approach about the focus group discussions and therefore meant those attending the focus groups were not truly representative of the main study population. This limits the transferability of the qualitative findings. The mean age of those attending the focus groups was much older than those recruited into phase one of the study, with mean age of 34 years (range 12-91) for phase one of the study and mean age 59 years (range 15-77) for the focus groups. Apart from one male aged 15
years and a female aged 37 years the remaining patients attending the focus groups were aged 62-77 years, meaning that the views of the predominantly younger population in phase one are not represented.

In addition, although fifty doctors were involved in phase one of the study none attended the focus group discussions, meaning that the findings represent the views of senior ENPs only. However, ENPs are routinely used to assess ankle injuries and their views and experiences are therefore valid and transferable.

5.4.2: Audit trail / quality control

There are several ways to enhance the trustworthiness of research findings, including the use of mixed research (or triangulation), respondent validation, transparency of study methodology and data analysis, and having two researchers independently analyse the data. Transcribing the audio recordings is the first stage of data analysis and enables the researcher to become immersed in the data (Holloway & Wheeler 2010), but I had no previous experience of transcribing and although this was attempted a decision to have the recordings professionally transcribed was taken to save time. I did however, listen to the audio recordings and check them against the transcriptions for accuracy before deleting them. In addition, as recommended by Greenhalgh and Taylor (1997) two researchers independently analysed the data from the transcripts using thematic analysis (Braun & Clarke 2006). The units of meaning were given initial codes and then grouped to form initial themes,
before being aggregated into the key themes for both the patient and the clinician focus group discussions. Using this framework a clear audit trail was established and enabled the researchers to discuss and compare the findings. Although I could have used the computer software package MaxQDA to assist with this I found using a pen and paper copy of the transcripts easier to work with.

Pope and Mays (2006) describe respondent validation as comparing the researcher’s account of the data with that of the study participants. Within this study the transcriptions and a list of the initial themes could have been sent back to individual patients to check for accuracy, which Lincoln and Guba (1999) regard as the strongest check on the credibility of qualitative research. Respondent validation can help to avoid misinterpretation or misunderstanding of the data (Sandelowski 1993) but can be difficult to achieve since it involves an interpretive judgement (Noland & Behi 1995; Lincoln & Gluba 1999). Within a focus group discussion it is often difficult for patients to remember exactly what they said and Kvale (1996) suggests that verbatim transcripts can be a source of embarrassment, if not distress, to patients since oral language, when transcribed, can appear confused, rambling and even incoherent. For the patient the repetitive nature of the transcript can be shocking and leave patients feeling they have been portrayed as somebody functioning at a lower level of intellectuality. As the researchers and participants play different roles in the research process it is imperative that the accounts of the discussion will be different (Sandelowski 1993). In addition, participants can be defensive or non-
critical due to the relationship with the researcher, and may also change their views over time (Holloway & Wheeler 2010). How this conflicting data is managed then causes its own problems for the analysis (Pope and Mays 2006). Due to the difficulty in recruiting to the focus groups, and the fact that a significant period of time had elapsed between the focus group occurring and the transcripts being prepared for analysis it was felt respondent validation would not enhance the research findings.

In order to improve interpretation of the findings the moderators at the focus group discussions could have recorded non-verbal interactions between the patients which could then be used to enhance understanding and interpretation during data analysis (Pope & Mays 2006). However, although an experienced facilitator was used to manage the focus groups discussions the research nurse and administrator acting as moderators were inexperienced and as such the notes they made during the focus group discussions did not contribute additional understanding to the data set.

5.4.3: Reflexivity

Reflexivity refers to the ways in which the researcher has shaped the data collected, including their preconceived ideas and experiences. As a lead nurse working with the ENPs and as the lead researcher for the study I did not attend any of the focus group discussions as I felt this may potentially bias the findings. Instead the focus groups were moderated by a senior lecturer who
was an experienced qualitative researcher who had no prior contact with the participants or any knowledge of the tuning fork test or the OARs, and therefore could not influence the discussion. Nevertheless, when analysing the data, I conducted a sensitivity analysis to ensure that I was mindful of my own pre-conceived ideas and assumptions regarding patient experiences within the ED (Sensitivity analysis – appendix J).

The focus group discussions provided a useful and practical method to gather wider data from multiple voices. Often discussion within the groups triggered a chain reaction from other participants to contribute innovative ideas and expose feelings which may not have been uncovered in interviews (Sim 1998). The inductive nature of a qualitative enquiry through focus group discussions revealed a rich insight into the experience and acceptability of the tuning fork test for patients and clinicians, and future research could use these findings to develop a postal and or telephone questionnaire that could be administered to compare and contrast a larger population to include the views of younger patients, doctors and junior ENPs on the use of the tuning fork test to detect ankle fractures.

5.5: Conclusion

The focus groups explored the experiences of patients and clinicians involved in the tuning fork test. The findings showed that patients and clinicians agreed on some issues such as the importance of information, the need to reduce the
number of x-rays taken, and the logistical difficulty in getting an appointment with GPs. Patients and clinicians also agreed on the fact that the tuning fork was not painful and the need to compare both ankles when performing the test. However, there were important areas in which they did not agree. The perception that patients come to the ED expecting an x-ray was not revealed by the patients in these focus groups but instead appears to be a pre-conceived expectation of the clinicians taking part. The patients also showed they are aware of the risks of repeated exposure to x-ray, irrespective of what the clinicians claimed.

This study set out to assess the diagnostic accuracy of the tuning fork test in the detection of ankle fractures and to assess its acceptability amongst clinicians and patients and the full study discussion follows in chapter 6.
Chapter 6 - DISCUSSION

The aim of this study was to assess the accuracy of the tuning fork test when used by multiple operators on patients already identified as OARs positive to the ‘malleolar zone’ and to explore the experiences of the patients and clinicians involved. As such a mixed methods study was the most appropriate methodology to use. This chapter includes the main discussion on the findings from both the tuning fork test and the focus group discussions. Triangulation of the data adds meaning to the figures of the diagnostic test and puts the patient and clinician views into perspective. This chapter contains an overview of the study methodology, the study population, a discussion of the main findings from both phases of the study, and the study limitations.

6.1: Overview of study methodology

6.1.1: Diagnostic test study

The original proposal included an ED, a minor injury unit and a walk-in-centre to ensure the findings could be transferable into the different emergency care settings. However, the walk-in-centre withdrew from the study before recruiting any patients in January 2010. Factors influencing their decision to withdraw were that the site specific investigation took six months to complete, the principal investigator went on long term sick leave, and in this this time the management structure of then unit changed. As a result the staff changed from being really keen to be involved in the study to having no interest in research at
all. The majority of the staff, including clinical senior leaders, could not see the benefit of undertaking research within their own care setting, and despite regular meetings to try and boost enthusiasm the walk-in-centre withdrew from the study without recruiting a single patient in January 2010. Although there were two other walk-in-centres in the same city neither had x-ray facilities.

Recruitment was already behind schedule at study sites A and B, and in anticipation of not meeting the target recruitment, the data collection period was extended by six months and the study opened to two mid-sized emergency departments (study sites C and D). Unfortunately this reduced the type of emergency care setting and patient population that the study findings could be transferable to, but were still applicable to the settings in which the tuning fork test would be used. The withdrawal of the walk-in-centre obviously had an impact on the rate of recruitment but clinical trials are notoriously difficult to conduct in the clinical setting due to work pressures, time constraints and lack of enthusiasm amongst staff members (Sitzia 2002) and all these issues were raised by the ENPs attending the focus group discussions.

All but one of the ENPs working at study sites A and B had been involved in clinical trials before and those attending the focus group discussions said that the tuning fork test was simple and straight forward to learn and use, and that the study had been ‘less faffy’ than other studies they had been involved in, yet they all reported ‘dropping off in winter’, and actively not recruiting patients.
when the department was busy. This therefore affected the consecutive recruitment of patients.

Study site A recruited the most patients (63%), whereas study site B recruited the larger proportion of their annual attendance figures when compared to the other study sites. This was possibly because study site B was a dedicated minor injury unit and as such ankle injuries would have made up a larger proportion of their attendances. In contrast the other three sites were all generic EDs attached to general district hospitals. The recruitment rate at study site D was disappointing but on review they had a limited ENP service and the principal investigator at the site initially restricted recruitment to the ENP team excluding doctors from recruiting patients into the study. Therefore when there was no ENP on duty no recruitment could take place.

However, on reflection Polit and Beck (2004) states that a researcher should aim to recruit 10% of the available patients into a clinical trial and when the number of patients attending study sites A and B were reviewed the number recruited at these two sites was equivalent to 13% (n = 1064) of the total ankle injuries attending (n = 8079) during the recruitment period. Therefore recruitment was not as slow as first anticipated and the target recruitment figure was reached and in fact exceeded (n =1358) within the allotted time limit.
6.1.2: Focus group discussions

Focus group discussions were chosen as the methodology to explore patient and clinician experiences of the tuning fork test. It was anticipated that this would allow the recovery of rich data about their lived experiences. However both the patient and clinician focus group discussions were poorly attended with only ten patients and ten clinicians attending. Since ankle injuries are deemed ‘minor’ injuries it may have been that coming to a focus group discussion to talk about their injury three months later would not have been seen as a priority by the patients. On refection it may have been better to undertake a short survey of the patient’s during their initial ED visit. A simple questionnaire designed to explore patient’s initial thoughts, the sensation felt, whether they attended ED with the expectation of receiving an x-ray and exploration of what rate of missed fractures they were prepared to accept may have captured more views soon after injury. In a similar way, clinician views could be sought after their involvement in the study.

6.2: Clinician experience

In contrast to Dissmann and Han (2006) this study did not restrict the number of clinicians involved in performing the tuning fork test and a total of 113 clinicians (60 nurses and 53 doctors) were involved in obtaining consent, performing the tuning fork test and collecting study data. All the nurses involved in assessing patients with the OARs were experienced emergency care nurses who had undergone post-registration training in the management of minor injuries and it was confirmed at the focus groups that all had a working knowledge of the
OARs prior to the start of the study. However the junior doctors, who were on two, three and six month rotations to the study sites, had rarely heard of the OARs prior to their involvement in the study. Unfortunately no doctors attended the focus group discussions and this study did not set out to assess the accuracy of the junior doctor’s ability to use the OARs but all received training on their use and were assessed in their application before they were permitted to recruit patients into the study.

6.3: Population & predictor variables

In total 1358 patients were recruited into the study but 45 were removed as they did not meet the study inclusion criteria. Surprisingly four of these were documented as having deformity to the ankle and as such required manipulation in the ED. Obvious deformity to the ankle excludes the use of the OARs as they were not developed to replace clinical judgement (Stiell et al 1992) and the recruitment of these patients identifies that clinicians may be inappropriately using the OARs for all ankle injuries. All the ENPs attending the focus groups had a working knowledge of the OARs and stated that ‘they just knew them’, but the inclusion of patients with deformity would indicate that they may not know them as well as they actually think, and it is therefore recommended that regular assessment is undertaken on any clinician using these rules to ensure they are applying them appropriately.

Only patient age and role of clinician performing the test were identified as clinically significant predictors of either an accurate or a positive tuning fork test
result. With adjusted OR of 1.021 (95% CI 1.014-1.029, p. <0.0001) and 0.993 (95% CI 0.985-1.000, p. 0.041) increasing age was a predictor of an accurate and a positive tuning fork test respectively. Nurses in this study were experienced in using the OARs, often assessed ankle injuries on a daily basis, and recruited more patients into the study. It is possible, therefore, that the nurses were more likely to record an accurate tuning fork test (adjusted OR of 1.595, 95% CI 1.174-2.167, p. 0.003). However, the nurses were less likely to record a positive tuning fork test when compared to doctors with an adjusted OR of 0.485 (95% CI 0.353-0.666, p. <0.001).

The reasons for this were not explored in the study but ENPs attending the focus groups had strong views that doctors ‘x-ray everything’ and have ‘lost the skills to assess minor injuries’ and this may reflect in the need for the doctors to record a positive tuning fork test to ensure the patient receives an x-ray if the test were introduced into clinical practice. The nurses recruited a mean of eighteen patients each, compared to only five for the doctors. Nurses, therefore, recruited more than three times more patients than the doctors did. However, one member in both the nurse and doctor groups recruited significantly more than the others. One ENP recruited 127 patients and one doctor, a clinical fellow at study site A, recruited 89 patients. When these two members are removed from the analysis the remaining nurses and doctors recruited a mean of fifteen and three patients respectively. Therefore, the difference between the nurses and the doctors may simply be that the more tuning fork tests a clinician performed the more confident they felt about the result and therefore the more accurate they became.
6.4: Inconclusive results

The OARs were used as a screening tool for inclusion into the study and as such only patients assessed as having bony tenderness to the malleolar zone were eligible to take part. However, 126 (10%) patients were recorded as being OARs ‘inconclusive’ and 39 patients were recorded as having an inconclusive tuning fork test.

Despite a review of the literature I could find no other studies assessing the effectiveness of the OARs or tuning fork test that include inconclusive results and it would therefore appear that this is the first study to report inconclusive results for both tests. All the patients recorded as inconclusive OARs went on to receive an x-ray. It was therefore assumed that if the tuning fork test was inconclusive these patients would also receive an x-ray if the test were introduced into clinical practice. Therefore patients recording inconclusive results were all treated the same and moved to the x-ray row of the 2x2 table for analysis. Unfortunately due to the timing of the focus groups, which took place before any data analysis had taken place, I was unable to explore why the clinicians recorded an inconclusive result or what constituted an inconclusive tuning fork test result and clinicians were not asked to explain their reasons for recording an inconclusive OARs result. However, 10% is a significant proportion of the sample and it is recommended that further studies attempt to define an inconclusive result, both for OARs and the tuning fork test.
6.5: Accuracy

The tuning fork test was accurate in 729/1313 patients, equating to a diagnostic accuracy of 56% (95% CI 53-58). The NPV was 98% (95% CI 94-97) meaning that a negative tuning fork test was very likely to result in a negative x-ray. However, there were 27 clinically significant fractures missed by the tuning fork test and the tuning fork test resulted in a sensitivity of only 84% (95% CI 78-89) and specificity of 51% (95% CI 48-54). This is lower than the sensitivity and specificity reported in the Dissmann and Han (2006) paper and lower that that expected at the start of the study. As outlined in section 3.3.1.1, although other measures of diagnostic accuracy are reported it is the sensitivity that is the most important measure of diagnostic accuracy in this study. Patients and clinicians will not accept a higher missed fracture rate than when the OARs are used in isolation and as such the tuning fork test does not improve the diagnostic accuracy of the OARs and the null hypothesis can be rejected, because the diagnostic accuracy is lower than when the OARs are used in conjunction with the tuning fork test. There were also a significant number of false positive tuning fork test, equating to 42% (n = 557) of the study population, which is much higher than the 24% reported in the Dissmann and Han (2006) study. The reason for this high false positive rate is unknown.

All the patients attending the focus group discussions recalled that they thought the test was accurate which is in contrast to the ENPs who all felt that the test was not accurate. Due to the timing of the focus groups no analysis had taken place before they were held and neither the moderator, nor the facilitator or the patients and clinicians knew the results of the study. Tuling (2000) states that
the vibrations from the tuning fork irritate the periosteal covering of the bone causing discomfort and pain in the presence of a fracture and prior to the start of the study, clinicians were advised that a sign of discomfort or pain when the tuning fork was applied indicated a positive tuning fork test. Clinicians were advised that discomfort or pain could be identified verbally or by the patient pulling their ankle away from the tuning fork, but this ‘response’ was left to individual clinicians to interpret. Patients were not informed what sensation they would feel when the tuning fork was applied as it was felt this would affect the accuracy of the test.

Both patients and ENPs attending the focus group discussions stated that the tuning fork test did not cause pain, and in fact both patients and clinicians recalled that the vibrations could be more on the good ankle than the injured one. Reviewing the data collection sheets some clinicians had written comments on the data collection sheets that the patient often reported feeling a decrease in sensation rather than pain when the tuning fork was applied to their injured ankle and this was also recorded as a positive test in some cases.

Interestingly, when the patients attending the focus groups were asked to recall what they thought their tuning fork test result had been three out of ten disagreed with the result recorded by the ENPs. In two of these the tuning fork test result recalled by the patient matched the result of the x-ray report. Unfortunately it was not within the remit of this study to have two clinicians perform the tuning fork test to confirm the test result as this would not be
standard practice if the test were introduced into clinical practice. However, it is recommended that this is included in future studies on the accuracy of the tuning fork test so that inter-operator reliability can be assessed.

6.6: Missed fractures

Diagnostic tests are rarely 100% accurate but any new test should be more accurate, cheaper, quicker and less invasive than the old test it replaces. The tuning fork test missed 65 fractures of which 38 were deemed not clinically significant, and were managed as sprains, and given no follow-up. None of these patients are known to have returned to the ED with the same injury. The definition for a clinically significant fracture was given at the start of the study and was the same as that given in the original work by Stiell et al (1992). However, unlike the work by Stiell et al (1992) this study also included children 12 years and over and Boutis et al (2001) and Gravel et al (2009) state that Salter-Harris I and Buckle fractures can also be deemed not clinically significant in children. However, this is in contrast to Clark and Tanner (2003) who acknowledge that although risks with these fractures are low all fractures in children have the potential to affect bone growth and all should therefore be treated as significant but the definition of a clinically significant fracture given at the start of the study was not changed.

Of the 27 clinically significant fractures missed by the tuning fork seven were managed as soft tissue injuries and discharged without follow-up. This is not the first study to report missed fractures with the tuning fork test. In addition to
the stress fractures missed by van Den Berg (2003) and Lesho (1997) Colwill and Berg (1958) and Bache and Cross (1984) missed pubic rami fractures and undisplaced hip fractures, which were also missed by Misurya et al (1987), and Moore (2009) missed avulsion fractures to the metatarsal and ankle, and buckle fractures to the clavicle.

6.7: Potential benefits

The results show that implementing the tuning fork test into clinical practice would result in a reduction in the number of ankle x-rays taken by as much as 47% and reduce the radiation exposure for patients, both of which were rated by patients as the main benefits of the study. In addition, although none of the patients attending the focus group discussions waited longer than three hours in ED they rated reducing waiting times as a benefit of the study. Patients assessed as tuning fork negative could have been discharged immediately after assessment, therefore reducing their length of stay by 67 minutes within this study, but this is at the expense of missing clinically significant fractures and the lower sensitivity reported is likely to be unacceptable to patients and clinicians.

6.8: Limitations of study

The final analysis was based on 1313 patient data sets and 172 fractures and as such the target recruitment was met for the diagnostic test study but the focus groups were poorly attended by both clinicians and patients with only ten of each group attending the focus group discussions. The age and gender of
the patients between the two phases of the study were different meaning that the two populations were not a true representative sample of each other. The patient focus groups were attended by a much older age group than those recruited into the diagnostic study, and all the ENPs attending had more than five years’ experience as an ENP. This therefore means that the findings from the focus groups may not be fully transferable across the whole study population and all clinician groups.

Variables such as age, gender, ethnicity, the degree of swelling, presence of distracting or previous injury, study site attended and role of clinician performing the test were deemed to be predictor variables for an accurate and a positive tuning fork test at the start of the study. Unfortunately the withdrawal of the walk-in-centre affected the mix of ethnicity in the population under study and only 26 participants were classified as being from a non-white ethnic group so the variable ethnicity had to be withdrawn from the analysis limiting the transferability of the results into non-white ethnic groups. The study also had a minimum age criterion of 12 years of age and as such the results are not readily transferable into younger populations.
Chapter 7 - CONCLUSION & IMPLICATIONS FOR PRACTICE

7.1: Conclusion

Ankle injuries account for approximately 1.75 million ED attendances annually, 8% of all minor injuries, and 10% of all x-rays requested within an ED (Wall 1999). Due to the high number of ankle x-rays requested in the early 1990s the OARs were devised by Stiell and colleagues (Stiell et al 1992, 1993) to try and improve clinician assessment of ankle injuries. Strengths of the OARs include

- They are one of the most validated clinical decision rules used in clinical practice and are widely used.
- They are easy to learn and apply
- They have been validated in adults and children
- They reduce the number of ankle x-rays taken in the majority of cases
- They have a consistently high sensitivity and the missed fracture rate is as little as 2%
- When compared to other ankle rules the OARs are consistently more sensitive

However, as with any clinical decision rules there are also weaknesses and these include
• They were devised for use on adult patients presenting with simple twisting and direct trauma to the ankle but clinicians regularly use them for all ankle injuries regardless of mechanism

• They were not designed to replace clinical judgement, but this study shows they are used even in the presence of obvious deformity by some clinicians

• Although easy to learn inexperienced clinicians often struggle to recite them in full

• They have a consistently low specificity in adults and children

• The prevalence of fractures in the population still receiving an x-ray is only 15%

• Caution is recommended when applied to children less than 6 years of age

• Loss of professional credibility in the event of missing fractures and the fear of litigation affect some clinicians use of the rules

Attempts to modify and improve the specificity of the OARs have led to a variety of results and leads to a decrease in the sensitivity in the majority of papers and this reduction is clearly not acceptable to clinicians or patients (Kerr et al 1994). X-rays are the traditional method for detecting fractures in the acute setting but they are costly and carry risks of exposure to radiation, albeit small, and this study set out to assess whether the accuracy of the OARs could be improved by using the adjunct of the tuning fork test.
Detecting fractures using sound waves was first described in 1823 (Peltier 1977), and Lippmann (1932), Colwill and Berg (1958), Peltier (1977), Bache and Cross (1984), Misurya et al (1987) File et al (1998), Moore (2009) and Tiru et al (2002) all describe using the technique to detect long bone fractures. However, the technique failed to detect ankle fractures in two studies (Colwill and Berg 1958; Moore 2009). Dissmann and Han (2006) were the first to focus solely on ankle injuries and the use of the tuning fork to detect fractures. The tuning fork test was used as an adjunct to the OARs and the results showed that x-rays could be reduced by 66% when used on patients assessed as having tenderness to the lateral malleolus. However, this study was small scale, omitted patients with swelling, and the tuning fork test was performed by a single operator. Although the results could not be readily transferred into other emergency care settings the findings were interesting and credible enough to further investigate the use of the tuning fork to detect ankle fractures.

This mixed methods study therefore set out to build upon the work of Dissmann and Han (2006). The main objectives of the study were to assess the accuracy of the tuning fork test and to identify whether there were any potential predictor variables to an accurate tuning fork test and to explore patient and clinician acceptance of the test. A two phase study was undertaken comprising a large diagnostic test study, which included multiple operators, multiple sites, a wide age group of patients and ankle injuries sustained by twisting mechanisms, and focus group discussions to explore patient and clinician acceptance of the test.
The final analysis included 1313 study participants which were a true reflection of the national population attending EDs with ankle injuries regarding age, gender, site of ankle injured and the prevalence of fractures in the population was 16%. This enables the findings to be readily transferred into other ED populations. However, the withdrawal of the walk-in-centre at the start of the study reduced the number of patients from different ethnic backgrounds to only 2% and caution should be taken when transferring the findings into non-white ethnic populations.

Accuracy of diagnostic studies are traditionally reported as sensitivity and specificity but these can often be misleading to clinicians so diagnostic accuracy and predictor values were also analysed.

The following is a list of what this study contributes to the current body of knowledge on the use of the tuning fork to detect fractures

- **Accuracy:** The tuning fork test lowers the accuracy of the OARs when used on lateral and medial ankle injuries

- **Missed fractures:** The tuning fork missed a number of significant fractures. Seven were managed as soft tissue injuries

- **Inconclusive results:** The study is the first to report inconclusive OAR and tuning fork test results and further research in this area is recommended.
- **Training**: Clinicians all received training prior to the start of the study but all reported not knowing if they were using the tuning fork correctly at the focus groups.

- **Predictor variables**: Operator and increasing patient age are predictors of an accurate tuning fork test.

- **Sensation felt and comparison with the ‘good’ ankle**: Patients attending the focus group discussions recalled that the tuning fork test did not cause pain, but described it as an intense vibration. Both patients and clinicians feel a comparison with the ‘good’ ankle is essential.

- **Acceptability**: The tuning fork test is accepted by patients and clinicians but is considered inaccurate by clinicians.

Each of these will be summarised below.

### 7.1.1: Accuracy

At 84% (95% CI 78-89) sensitivity and therefore the diagnostic accuracy is lower than when the OARs are used in isolation and the number of missed x-rays in this study will be unacceptable to patients and clinicians.

### 7.1.2: Missed fractures

There were a total of 27 clinically significant fractures missed; however, eight of these were managed as soft tissue injuries. These were
• small avulsion fractures between 3 and 4.5mm in depth
• a buckle fracture to the tibia
• a patient presenting 19 days post injury and
• a patient presenting 3½ weeks post injury

It could be argued that these fractures can also be deemed not clinically significant as none of the patients received treatment for a fracture, none were given follow-up, and none returned to the ED, but the definition of a clinically significant fracture had been set at the beginning of the study and I was not going to change this as this could affect the credibility of the findings.

Distal fibula fractures were frequently missed (n = 10) by the tuning fork and accounted for 37% of the missed fractures in the study. Review of the ED records of eight of these found that the site of maximal tenderness did not match the site of the fracture and clinicians may have incorrectly placed the tuning fork in these patients. The reasons for this are unknown but may be due to the fact that the associated soft tissue injury that accompanies these fractures may distract the pain away from the fracture site. This has implications for not only the implementation of the tuning fork test but also the accuracy of the assessment of bone tenderness by the clinicians. This leads to the need to undertake further work into the use of the tuning fork and a recommendation that inter-operator reliability is assessed, with two clinicians performing and interpreting the result in the presence of a negative test.
7.1.3: Inconclusive results

This is the first study to report inconclusive OAR and tuning fork test results. Unfortunately the study did not explore what an inconclusive OAR result meant to clinicians but 10% (126) of the study population is a significant number. All these patients went on to receive an x-ray as per study protocol. In addition, 39 patients were reported as having an inconclusive tuning fork test, and for analysis it was assumed that these patients would also go on to receive an x-ray if the tuning fork test were introduced into clinical practice in the future. It is recommended that further studies should aim to define clinician interpretation of an inconclusive OAR and tuning fork test result.

7.1.4: Training

This is the largest study to date on the use of the tuning fork to assess fractures, not only because of the number of patients recruited (N = 1358) but also because of the number of clinicians involved performing the tuning fork test (N = 113). Although this was a clinical trial it was important to ensure that the situation in which the tuning fork was used was kept as true to normal clinical practice as possible. Therefore clinicians were trained in the use of the tuning fork test before their involvement in the study and then left to perform the test independently. It became evident throughout the analysis of the focus group discussions that clinicians had doubts as to whether they were using the tuning fork correctly and, although additional training was available, only one ENP requested this. This has implications for the implementation of the tuning fork test and identifies that on-going support and reinforcement on the use of the
tuning fork to detect fractures is needed. The removal of four post consent patients for having ‘obvious deformity’ identifies that clinicians may be using the OARs on all ankle injuries, regardless of mechanism, when they were in fact designed for use on simple, twisting and direct blows only. They were not designed to replace clinical judgement.

7.1.5: Predictors of an accurate tuning fork test

Increasing patient age and having a nurse perform the test were predictors of an accurate tuning fork test. However, this may simply be because the nurses recruited a significantly larger proportion of the study population than nurses and as such became more adept at performing the tuning fork test. Although it was anticipated that swelling could affect the tuning fork test an accurate tuning fork test was in fact more likely with significant swelling compared to no swelling. This finding should be treated with caution as the number presenting with swelling was much lower than those identified as having significant swelling, 38 compared to 152 respectively.

7.1.6: Sensation felt and comparison with the ‘good’ ankle

Previous studies state that the tuning fork causes pain as it irritates the bone periosteum in the presence of a fracture (Tuling 2000). However, none of the patients returning to the focus groups recalled the tuning fork test was painful, instead describing the sensation as an intense vibration or tingling. In addition, the ENPs recalled that in some patients the vibrations felt were less on the
injured ankle than the good ankle, and both patients and clinicians suggested that a comparison with the ‘good’ ankle is a necessity. This therefore has implications when implementing the test into clinical practice as there has to be a clear indication of what constitutes a positive tuning fork test.

7.1.7: Acceptability

Attendance at the focus groups was low with a significant number of patients declining to take part. Neither patients nor clinicians were asked why they did not want to take part, but a few patients offered that they didn’t have time in their busy work life schedules to return to talk about such a minor injury. Unfortunately the population attending the focus groups did not truly reflect the main study population and so the findings need to be treated with caution. However, the data obtained from the focus groups was relevant, focused, gave an in-depth view of clinician and patients views on the tuning fork and therefore met the aims of the study.

Patients and clinicians disagreed when it came to assessing accuracy of the tuning fork test, with patients stating that the test was accurate and clinicians stating that it was not. Despite this both clinicians and patients accepted the tuning fork test as a method to assess ankle injuries.

The focus groups provided sufficient data to devise a questionnaire to explore further patient and clinician acceptance of the tuning fork test. However, on
reflection a short patient survey, undertaken at the time of the patient’s initial ED visit would have captured the views of more patients and may have produced more data.

### 7.2: Implications to practice & recommendations for future research

Although training on the use of the OARs is given during ENP training this study highlights the need to provide regular updates to ensure the rules are being adhered to and not deviated from. Junior doctors are not routinely taught the OARs prior to their ED placements and may need to be assessed in their implementation of the rules before being allowed to independently assess patients presenting to ED with ankle injuries. This in itself may reduce the number of ankle x-ray requested. Providing patients with adequate information appears to be the key to gaining patient compliance with the injury and it would appear that the perception a patient wants an x-ray may not exist, except in the minds of the clinicians.

Using the tuning fork as an adjunct to the OARs appears to be acceptable to patients and clinicians and could result in a reduction in x-rays, but does not improve the diagnostic accuracy of the OARs and misses some clinically significant fractures. It is anticipated that the lower sensitivity will not be accepted by clinicians but in where the tuning fork test is implemented into clinical practice it is recommended that two operators perform the tuning fork test when a negative result is obtained to ensure inter-operator reliability.
Further studies need to focus on the use of the tuning fork in the detection of fractures not only to the ankle but also other extremities. Future studies need to include multiple operators of varying experience, assess inter-operator reliability, and explore the phenomenon of inconclusive OAR and tuning fork tests. The focus group discussions gave an insight into patient and clinician experience of the tuning fork but the findings represent the older population and senior ENPs only. It is recommended therefore that a questionnaire is developed; using the data obtained from the focus group discussions, to compare and contrast a larger population on the acceptability of the tuning fork test in the detection of fractures.
Chapter 8 – REFLECTION

This chapter includes a reflection on my journey through the Professional Doctorate training, my role as a practitioner-researcher and the future.

8.1: On undertaking Professional Doctorate training

When I started out on the Professional Doctorate I had just completed my Masters (MSc) degree and was about to start a consultant practitioner training program with NHS South Central. The aim of this training program was to prepare me for the role of a consultant nurse and gave me the benefit of protected study time. As described by Batchelor and Napoli (2006) I felt passionate, motivated and excited about the things to come.

A colleague had started his Professional Doctorate twelve months earlier whilst another was due to start the Doctor of Philosophy (PhD) and in keeping with the work of Ellis (2006) I was frequently asked why I had chosen the Professional Doctorate over the more tradition PhD. Nursing colleagues could not grasp the notion that I would be a ‘Dr’ on completion of the program of study and did not know the difference between the two degrees, whilst medical staff thought the Professional Doctorate was a “watered down PhD” (Dreher 2011, p. 403). I initially found it difficult to explain the differences to colleagues, but I knew the requirement to undertake the training at Portsmouth emphasized eligibility criteria as “experienced practitioner”, and I wanted to access continuous
training, assessment and peer support through attending taught study sessions with other clinicians. However, more than that, the difference was unclear; I was to be honest confused myself. I then found the definition given by Taylor (2007) which simply states that a Professional Doctorate produces a ‘researching professional’, and a PhD a ‘professional researcher’!

I knew then I was doing the right doctorate pathway. I wanted to be a researching professional and already had experience of undertaking research in the clinical setting (Welling 2006; Welling 2007). I had a good track record of setting my own learning objectives and although I knew I had the motivation to study in isolation previous reflection shows that I learn best when I am in a group situation and I enjoyed the taught element of the programme.

When starting out little did I know how big this current study would become! Batchelor and Napoli (2006) define the journey of the Doctoral student as

“taking ship and watching solid land drift away”................. “you expect to make steady progress and move forward. But sometimes the exact opposite happens: you become as becalmed as the fleet at Aulis, with no breath of wind to move you in any direction”..............“when the wind changes and you do start to move again”............”you are pulled in different and sometimes conflicting directions by the currents” (p. 13-14).

This metaphor of ebb and flow accurately describes my feelings throughout the Professional Doctorate. As I set out on my journey I already knew the topic I
wanted to research, I had done some initial searching for what was already known on the topic and had already discussed the topic with work colleagues to gauge their opinion and support but I never anticipated that there could be so many barriers to success. Applying for funding added a year or more to the project, delayed the application for ethical approval, and significantly changed the size and scope of the project undertaken. Whilst waiting for the funding decision I often felt I was ‘up the creek without a paddle’ and the project seemed to stall. In keeping with Batchelor and Napoli (2006) I too became anxious and frustrated at not being able to progress and during this time I found the support I received from my peers invaluable. Meeting up with the study group regularly helped to keep my mind focused on the study as identified by (Lee 2008). When the RfPB peer reviews were received, the project accepted subject to achievable amendments, it gave the boost to my enthusiasm and confidence which I needed to continue with the project. I knew then that I had a project that could make a big difference to patients and the NHS not only locally but also nationally.

Although the application for funding delayed the study by a year, on reflection, without the funding, I could not have completed the study. The funding gave me the opportunity to employ a research nurse and a support worker to assist with the day-to-day administration of the study, and ensured that all the study costs were covered. On receiving the funding I was naïve and did not realise that I would have financial and written reports to complete as well as completing the portfolio accruals database. This was in addition to the participant data that had to be entered onto the Edge™ database used by the Trust in which I worked.
These two databases were not linked and were not compatible with transference of data so each patient demographic had to be entered twice. This took time and without the assistance of the research nurse and support worker there is no way I could have kept this up to date whilst working full time and leading the research across the study sites.

Neither the research nurse nor the support worker had previous research experience so I had to supervise and train them in their roles, which enabled me to improve my skills in management and leadership. I was able to help them identify skills gaps where their own knowledge was lacking and help them identify potential training opportunities to help shape their own research opportunities. By the end of the study the research nurse was sufficiently experienced to continue in her role with the Research and Development team at the hospital in which I work and is now the lead research nurse for two further studies. She has developed the same passion for research as I have; of which I am immensely pleased.

Ethics caused another time where emotions could again be likened to that of a stormy ride at sea, and at this time I really did consider quitting altogether. I had withheld applying for ethics whilst waiting for the funding decision but two weeks before I was due to submit the ethical application process transferred to IRAS. The original application form was found to be irretrievable and the whole application process had to be started again. On the day of the ethical review meeting I heard the news that my step-son had tragically died and this brought
home the personal sacrifices required when committed to undertake such an intense program of study.

Once the study had started recruitment was slower than expected and the withdrawal of the walk-in-centre again evoked a mix of emotions. I was angry because they had withdrawn but also felt sorry for the staff working there as they were stuck in a culture where research and evidence based practice did not appear to play a part. However, I was also relieved that the study could be opened up to two other departments who were keen to get involved, but had missed out on the opportunity at the beginning of the study.

My experience of conducting a randomised controlled trial on burns (for my master’s research project) (Welling 2007) had already identified that work pressures and lack of time were potential barriers to recruitment in clinical trials but I still found myself becoming frustrated at the lack of recruitment by some ENPs. At times I felt the study was running away from me and was going to fail simply because of a lack of enthusiasm by others. ‘Why aren’t they as keen as me?’ was a question I often asked myself. I could not possibly be there all the time recruiting every single patient. The art, and it is an art, of keeping the study focused in everyone’s minds was challenging. Posters, pens, donuts, and emails were techniques employed to boost recruitment and I was delighted once sufficient numbers had been recruited and analysis could start.
8.2: On my role as a practitioner-researcher

I have been working in the role of consultant nurse since 2008. This role focuses on the domains of expert clinical practice, research, service development and education. It is often difficult to juggle all these domains satisfactorily and 75% of my work is focused on delivering clinical care to patients with the remainder split between the other three domains. However, research plays a big role in my work and as stated previously I have now completed three of my own studies. My focus has always been on quantitative methods so the suggestion at the beginning that this study would also include qualitative filled me with dread. However, I really enjoyed this element of the study and I would not have learnt these skills if I had not completed the relevant doctorate training. Learning how to analyse the transcripts has increased my ability to articulate my own feelings, improved my listening skills, and helps me synthesis narrative data.

I am naturally inquisitive which is an essential attribute of a practitioner-researcher (Dadds n.d.) and I regular challenge current practice. The barriers encountered during the progress of the study have made me reflect on practice and given me invaluable insight into the pitfalls of mixing research and clinical practice. However, it has also made me more determined to succeed. I do however, become frustrated at the barriers that exist for nurses working in advanced roles and the slow pace of change with those who regulate research practice. As a nurse I am able to be the principal or lead investigator for a study but according to the Good Clinical Practice Guidelines in the Declaration of Helsinki 2008 the clinical care of patients in a study still has to be the
responsibility of a medical physician, and this will not change with the title Dr before my name. This is despite the fact that as a consultant nurse I am expected to work autonomously, manage clinics and am often the most senior clinician working within the minors or paediatric areas of the ED.

### 8.3: On the future

Working within the emergency department has always meant having to work under pressure and be receptive to change and completing the Professional Doctorate training has increased my confidence, and ensured I am better equipped to articulate my thoughts. In keeping with Sanders, Kuit, Smith, Fulton and Curtis (2011) I have developed the skills I need to reflect on practice, and critically evaluate and question current practice. As an introvert I used to detest having to attend meetings but the training of the Professional Doctorate, chairing the Study Steering Group, and the dissemination of the study at conferences, has given me the confidence I need to freely articulate my thoughts when appropriate and increased my leadership skills to support junior staff in fulfilling their goals.

Research is now allocated within my job plan and I am currently the principal investigator at the trust for three national studies. I will continue in my role of consultant nurse and as such continue to develop and initiate other research activities within the department in which I work whilst maintaining a clinical focus. My ultimate goal is to become a research fellow, but only if I can maintain my clinical focus. Undertaking the Doctorate training is unlikely,
particularly in the current economic crisis, result in financial reward or promotion
but will, I hope, enhance my professional credibility amongst fellow colleagues
(Ellis 2006) nationally as well as locally.

To conclude, the words below sum up, I feel, all that I anticipate my future after
successful completion of the Professional Doctorate training should be –

Practitioner research is a challenging and exacting undertaking. It
demands open-mindedness, courage in the face of self-critique and public
sharing, emotional fortitude in dealing with uncertainty and profound
change, spiritual energy in sustaining curiosity, compassion and the
eternal search for new, improved practices. Those who take this journey
demonstrate the highest levels of professionalism in using and
developing these qualities. They offer innovative role models for others
and help build a new climate for responsible and accountable
professionalism
Dadds (n.d., p. 9).
Chapter 9 - REFERENCES


Sim. (1998). Collecting and analysing qualitative data: issues raised by focus groups. Journal Advanced Nursing, 28(2), 345-352


Sitzia, J. (2002). Barriers to research utilisation: The clinical setting and nurses themselves. Intensive and Critical Care Nursing, 18(4), 230-243


Chapter 10 - APPENDICES

Contents

A – Patient information sheet
B – Age appropriate child and parent information sheet
C – Consent form for phase one of the study
D – Data collection sheet
E – Consent for patient focus group discussions
F – Patient invitation letter to attend focus group discussions
G – Patient interview schedule
H – Consent form for clinician focus group discussions
I – Interview schedule for clinician focus group discussions
J – Sensitivity analysis
Appendix A

Patient information sheet
Has anyone else reviewed?
This study has been reviewed by the (NAME OF LREC HERE) and an ethics committee from the University of Portsmouth.

Who is organising and funding the research?
Anne Welling is the lead researcher and is a consultant nurse working in the emergency department at the (NAME OF NHS TRUST HERE). She is studying for a Professional Doctorate degree at the University of Portsmouth. Funding for the study has been received from the Research for Patient Benefit Scheme and the project is sponsored by Portsmouth Hospitals NHS Trust.

Contact details
If you require more information or want to discuss the project you can contact the research team on the numbers below:

Lead Researcher – Anne Welling
*Tel: TBC
*Email: TBC
*Or by post at the following address, TBC

Project Research Nurse – TBC
*Tel: TBC
*Mobile: TBC
*Email: TBC

Lead Nurse at this site – TBC
Tel: TBC
Email: TBC

Project Support Worker – TBC
Tel: TBC
Email - TBC
Why are we doing research?
Research is a way to find out the answers to questions. The question we want to answer is “Can a tuning fork be used to assess for a broken bone after an ankle injury”.

Who will be invited to take part in this project?
If you are attending the department because you have twisted your ankle you will be invited to take part in this project but only if you are assessed as having tenderness to one or more of the bones in your ankles and need an x-ray.

What is being tested?
A tuning fork vibrates gently and transmits sound waves through bone. It is usually used to assess hearing but we want to see if it can identify whether a bone has been broken after injury to the ankle. The result of the tuning fork test will be compared to the x-rays to see how many times the results agree and disagree. This will show us if the tuning fork is an accurate way of assessing an ankle injury.

Do I have to take part?
No, it is up to you. The nurse or doctor who has assessed your injury will ask you to sign a form giving your consent. You will be given a copy of this information sheet and the signed consent form to keep. You are free to stop taking part in the project at any time without giving reason, just let the nurse or doctor know. It will not affect the care you receive.

What will happen to me if I take part?
This project has two parts.

Firstly, the doctor or nurse will place a tuning fork on both your ankles and record your response. You will then be sent for an x-ray of your ankle and the injury will be managed according to the diagnosis.

We need 1300 patients to complete this part of the project.

For part 2 of the project we want to hold a series of meetings to find out what patients thought of the tuning fork test. These meetings are called ‘Focus Groups’ and conversations that take place are recorded. Each meeting will last about 1-2 hours and take place in a local hotel. Refreshments will be provided. If you consent to us contacting you about taking part in a meeting we will do so within 3 months of your visit today. If you attend a meeting we will give you a £10 book token, and pay you reasonable travel expenses (not taxi fares).

We need 60 patients to complete this section of the project.

Are there any risks involved if I take part?
The risks from x-rays of the ankle are negligible. If you take part the dose and number of x-rays you receive will be the same you would get if you did not take part in the project. There is no known risk from the tuning fork test, but it can sometimes be a little bit uncomfortable. If you are pregnant and the benefit of x-ray is deemed to outweigh the risks to the unborn child you can still take part in the project if you want to as the tuning fork test is not known to be harmful in pregnancy.

What are the possible benefits of taking part?
We are hoping that the tuning fork test will help us improve how we assess ankle injuries in the future. This will result in a reduction in the number of ankle x-rays we take, and therefore reduce the cost involved. It is anticipated that patients will spend less time in the department as they will not need to wait for unnecessary x-rays.

Will you inform anyone else of my inclusion in the project?
No, we do not need to inform your GP or any other specialist as the project will not interfere with any other treatment you may be receiving. All your medical details will be kept private if you take part. No personal data will be kept for the purpose of the study as all the project data will be entered on a specially designed data collection sheet.
Appendix B

Age appropriate child and parent information sheet
and the x-ray.

Has anyone else checked the study is OK to do?
Before any research can happen it has to be checked by a group of people called a Research Ethics Committee. They make sure that the research is fair. This project has been checked by (NAME OF LREC HERE) and an ethics committee from the University of Portsmouth.

Who is organising and funding the research?
Anne Welling is the lead researcher and is a consultant nurse working in the emergency department at (NAME OF NHS TRUST HERE). She is studying for a degree at the University of Portsmouth. Funding has been received from the Research for Patient Benefit Scheme and the project is sponsored by Portsmouth Hospitals NHS Trust.

Contact details
If you require more information or want to discuss the project you can contact the research team on the numbers below:

Lead Researcher – Anne Welling
*Tel: TBC
*Email: TBC
*Or by post at the following address, TBC

Project Research Nurse – TBC
*Tel: TBC
*Mobile: TBC
*Email: TBC

Lead Nurse at this site – TBC
Email: TBC

Project Support Worker – TBC
Email - TBC
Why are we doing research?
Research is a way to find out the answers to questions. The question we want to answer is “Can a tuning fork be used to assess for a broken bone after an ankle injury”.

Who will be invited to take part?
You will be invited to take part if you have twisted your ankle and are aged 12 years or older, but only if you are assessed as having pain to one or more bones of your ankles and need an x-ray.

What is being tested?
When tapped against a hand a tuning fork vibrates gently and makes a quiet sound. When the tuning fork is placed over bone it sends the sound waves down the bone. A tuning fork is usually used to assess hearing loss but we want to see if it can identify if there is a break in a bone. The result of the tuning fork test will be compared to the x-ray result to see how many times they agree and disagree.

Do I have to take part?
No, it is up to you. The nurse or doctor will ask you to sign a form giving your consent. Your mum or dad will also need to sign the consent form if you are aged 12 to 15 years. If you want to take part in this part of the project your mum or dad will also need to attend. If you allow us to contact you about taking part in a meeting we will do so within 3 months of your visit today. If you attend a meeting we will give you a £10 book token. We need 60 patients to complete this section of the project.

What will happen to me if I take part?
This project has 2 parts. You can take part in the first part of the project without taking part in the second part.

First, the doctor or nurse will place a tuning fork on both your ankles and record your response. You will then be sent for an x-ray of your ankle and the injury will be managed according to the diagnosis. We need 1300 patients to take part in this part of the project.

For part 2 of the project we want to hold some meetings to find out what patients think of the tuning fork test. These meetings are called ‘focus groups’ and any talk that takes place will be recorded. Each meeting will take about 1-2 hours in a local hotel. Drinks and biscuits will be provided. If you want to take part in this part of the project your mum or dad will also need to attend. If you allow us to contact you about taking part in a meeting we will do so within 3 months of your visit today. If you attend a meeting we will give you a £10 book token. We need 60 patients to complete this section of the project.

Is there anything I should be worried about if I take part?
The risk from x-rays of the ankle is minimal. If you take part in the study the dose and number of –rays you receive will be the same you would get if you did not take part in the project. There is no known risk from the tuning fork test, but it may be a little bit sore.

What are we trying to make better?
We are trying to improve how we assess ankle injuries in the future. We want to try and reduce the number of x-rays we take. This will result in patients having to spend less time in the department and will reduce the amount of money we spend on x-rays.

Will anyone else know I am doing this?
Your parents will need to know if you are less than 16 years old. We will not need to inform your GP. All your medical details will be kept private if you take part. The only information we keep will be on a specially designed sheet which will allow us to record the result of the tuning fork test.
Child may be receiving. All your child’s medical details will be kept private if they take part. No personal data will be kept for the purpose of the study as all the project data will be entered on a specially designed data collection sheet.

Has anyone else reviewed the study?
This study has been reviewed by the (LREC NAME HERE) and an ethics committee at the University of Portsmouth.

Who is organising and funding the research?
Anne Welling is the lead researcher and is a consultant nurse working in the emergency department at the (NAME OF NHS TRUST HERE). She is studying for a Professional Doctorate degree with the University of Portsmouth. Funding for the study has been received from the Research for Patient Benefit Scheme and the project is sponsored by Portsmouth Hospitals NHS Trust.

Contact details
If you require more information or want to discuss the project you can contact the research team on the numbers below:

Lead Researcher – Anne Welling
*Tel: TBC
*Email: TBC
*Or by post at the following address, TBC

Project Research Nurse – TBC
*Tel: TBC
*Mobile: TBC
*Email: TBC

Lead Nurse at this site – TBC
Email: TBC

Project Support Worker – TBC
Email - TBC

The Ankle Injury Tuning Fork Research project
Parent Information Sheet

If your child is aged 12 years or over and attending because they have injured their ankle they may be asked to take part in a research project. Your child will have been given an information sheet to read. Please read this parent information sheet which will answer some of the questions you may have. If you child wants to take part in the project we will need your consent if the child is aged 12 to 15 years.

TRUST LOGO HERE

Version 01/dated 1st May 09

252
Why are we doing research?
Research is a way to find out the answers to questions. The question we want to answer is “Can a tuning fork be used to assess for a broken bone after an ankle injury”.

Who will be invited to take part in this project?
If your child is attending the department with a twisting ankle injury they will be invited to take part in this project but only if they are assessed as having tenderness to one or more of the bones in their ankle and need to have an x-ray.

What are we testing?
A tuning fork vibrates gently and transmits sound waves through bone. It is usually used to assess hearing but we want to see if it can identify whether a bone has been broken after an injury to the ankle. The result of the tuning fork test will be compared to the x-ray to see how many times they agree and disagree. This will show us if the tuning fork is an accurate way of assessing an ankle injury.

Does my child have to take part?
No, it is up to your child. However, if your child is aged 12-15 years you will also need to give your consent for them to take part in the project. If you are happy for your child to take part the nurse or doctor will ask you and your child to sign a form giving your consent. You can keep this information sheet and a copy of the signed consent form to keep. You and your child are free to stop taking part in the project at any time without giving reason, just let the nurse or doctor know. It will not affect the care your child receives.

What will happen to my child if they take part?
This project has two parts. Your child can take part in part 1 of the project without taking part in part 2.

Firstly the doctor or nurse will place a tuning fork on your child’s ankles and record their response. They will then be sent for an ankle x-ray and the injury will be managed according to the diagnosis.

We need 1300 patients to take complete this part of the project.

For part 2 of the project we want to hold a series of meetings to find out what patients and parents think of the tuning fork test. These meetings are called ‘Focus Groups’ and conversations that takes place are recorded. Each meeting will last about 1-2 hours and take place in a local hotel. Refreshments will be provided. If you consent to us contacting you and your child about taking part in one of these meetings we will do so within 3 months of your visit today. If your child attends a meeting we will give them a £10 book token, and pay you reasonable travel expenses, but not taxi fares.

We need 60 patients to complete this section of the project.

Are there any risks to my child if they take part?
The risk from x-rays of the ankle are negligible. The x-rays your child received will be the same as they would receive if they did not take part in the project. There is no known risk from the tuning fork test, but it can sometimes be a little bit uncomfortable.

What are the possible benefits of allowing my child to take part?
We are hoping that the tuning fork test will help us improve how we assess ankle injuries in the future. This will result in a reduction in the number of ankle x-rays we take, and therefore reduce the cost involved. It is anticipated that patients will spend less time in the department as they will not need to wait for unnecessary x-rays.

Will you inform anyone else of my child’s inclusion in the project?
No, we do not need to inform the GP or any other specialist as the project will not interfere with any other treatments your
Appendix C

Consent form for phase one of the study tuning fork study
Consent Form - Tuning fork study

Patient study number

- I confirm that I (patient)……………………………….. ...(and my parent/carer where applicable) have read the information sheet (version 01/Aug08) and have had the opportunity to ask questions.

- I (and my parent/carer where applicable) understand that participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.

- I (and my parent/carer where applicable) understand that sections of the medical notes may be looked at by responsible individuals from regulatory authorities where it is relevant to taking part in the research.

- I (and my parent/carer where applicable) give permission for these individuals to access my records

- I (and my parent/carer where applicable) agree to take part in the above study.

- **Do you consent to being contacted to take part in a focus group within 6-months?**  YES / NO
  (in the event of a child <16years both the parent and the child must agree before choosing Yes)

Name of patient:…………………………………………………………  ………… Date:………………

Signature:………………………………………………………………………………………………

Name of parent/carer (if applicable)……………………………………………………………..Date:………………

Signature:………………………………………………………………………………………………

Name of person taking consent:……………………………………………….….Date:………………

Signature:………………………………………………………………………………………………

Give the patient the original consent form after photocopying and putting a copy in the study folder

Trust logo here
Appendix D

Data collection sheets
Tuning fork study - Data collection form:

**ENP / nurse assessing patient to complete part A of this form**

**Part A:**

Date............

Age..................Gender: M / F

Ethnic origin: White / Black / Asian / Chinese / Mixed

Injured ankle: Right / Left

Bony tenderness: Lateral Malleolus / Medial Malleolus / distal fibula

Are the Ottawa Rules Positive / Inconclusive

Degree of swelling: None + ++ +++

Is the lateral and / or medial malleoli bony prominence visible? YES / NO

Are there any distracting injuries: Yes / No (please specify)............................

Is there any previous injury at site: Yes / No (please specify)............................

Tuning fork test: UNINJURED ankle:

Does the patient report or shows signs of discomfort on application: Yes / Unsure / No

Is the response a withdrawal from pain / verbal response / other (please specify)

INJURED ankle:

Does the patient report or shows signs of discomfort when the tuning fork test is applied to
- The site of maximal tenderness? Yes / Unsure / No
- 6cm proximal? Yes / Unsure / No:

Is the respond a withdrawal from pain / verbal response / other (please specify)

Name of person performing tuning fork test:.................................
**Part B:**
**To be completed by chief investigator or principal investigator only**

<table>
<thead>
<tr>
<th>Study number:</th>
<th>Study site: A / B / C / D</th>
</tr>
</thead>
</table>

**Time of arrival:**
**Time of assessment:**
**Time of x-ray (taken from PACs):**
**Discharge time:**

**X-ray result:** *(transcribed from radiologist report)*

**Result of tuning fork test:**
Positive / negative / inconclusive

**Data input onto computer by:** ............................................
Tuning fork study - Data collection form:

**ENP / nurse assessing patient to complete part A of this form**

**Part A:**

**Date:**

**Age:**

**Gender:** M / F

**Ethnic origin:** White / Black / Asian / Chinese / Mixed

**Injured ankle:** Right / Left

**Bony tenderness:** Lateral Malleolus / Medial Malleolus / distal fibula

**Are the Ottawa Rules:** Positive or Inconclusive

**Degree of swelling:** None + ++ +++

**Is the lateral and / or medial malleoli bony prominence visible?** YES / NO

**Are there any distracting injuries:** Yes / No (please specify)..................

**Is there any previous injury at site:** Yes / No (please specify)..................

**Tuning fork test:**

**INJURED ankle:**

Does the patient report or shows signs of discomfort when the tuning fork is applied to:
- The site of maximal tenderness? Yes / Unsure / No
- 6cm proximal Yes / Unsure / No

Is the response a withdrawal from pain / verbal response / other (please specify)

**UNINJURED ankle:**

Does the patient report or shows signs of discomfort on application:

Yes / Unsure / No

Is the respond a withdrawal from pain / verbal response / other (please specify)

**Name of person performing tuning fork test:**.........................
Part B:
**To be completed by chief investigator or principal investigator only**

Study number:  
Study site: A / B / C / D

Time of arrival:  
Time of assessment:  
Time of x-ray (taken from PACs):  
Discharge time:

X-ray result: *(transcribed from radiologist report)*

Result of tuning fork test:
Positive / negative / inconclusive

Data input onto computer by: ..............................................
Appendix E

Consent form for patient focus group discussions
Informed consent form:

Tuning fork study focus group participation

Thank you for agreeing to be involved in part two of the tuning fork study. This form outlines the purpose of this part of the study, describes your involvement and your rights as a participant.

The purpose of this study is:
- To gain insight into whether patients accept the use of the tuning fork test to assess the need for x-ray following an ankle injury, and whether they consider reduced waiting times or receiving an x-ray as the most important factor when they seek medical attention with an ankle injury.

The benefits of the research will be:
- To increase our understanding into what patients rate as the most important factor when they attend emergency care environments with a minor injury.
- Use information obtained to plan new ways of working that are accepted by patients and increase satisfaction with the service.

The method that will be used to meet this purpose is a series of focus groups. A focus group allows patients involved in the first part of the study to get together and discuss their thoughts about the tuning fork test. The focus group is being led by (NAME OF FACILITATOR HERE) who is a senior lecturer at the University. Her role is to make sure that every member of the group has a chance to share their views. The facilitator will start the conversation within the group by asking some open questions which should then encourage a discussion to take place.

The discussion will be audio recorded to help us accurately capture the thoughts of those present in their own words. The recordings will only be heard by the members of the research team for the purpose of transcribing. Once recordings have been transcribed they will be deleted. Your direct quotes and those of others within the group will be used to write a final report. However, your name and other identifying information will be kept anonymous.

You have the right to withdraw from this part of the study at any time. In the event that you do withdraw from the study your recorded views will not be used in the final report of the study. However, the data collected from the tuning fork test will continue to be used.

By signing this consent form I certify that I ……………………..agree to the terms of this agreement. (print full name here)

Signature……………………………………………..Date…………………………
Appendix F

Patient invitation letter to attend the focus group discussions
Date:

Dear:

You recently participated in the Ankle Injury Tuning Fork study at (NAME OF STUDY SITE HERE). At that time you consented for us to contact you about taking part in the second part of the study – a focus group.

Just to remind you a focus group is where we invite a number of patients who participated in the study to meet and discuss their experiences of receiving the tuning fork to assess their ankle injury. The meeting will be led by (NAME OF FACILITATOR HERE) who is a lecturer at (NAME OF UNIVERSITY HERE). The meeting is expected to take between one and two hours. The conversation will be recorded on a digital recorder. Light refreshments will be provided free of charge. In addition, you will receive a ten pound voucher as a gesture of goodwill and reasonable travel expenses can be claimed for. However, taxi fares will not be paid.

We would like to invite you to one of the forthcoming focus groups. A copy of the patient information sheet is enclosed for your information.

Please complete the enclosed form and return to the Admin Support Worker in the pre-paid envelope provided.

Thanking you in anticipation

Anne Welling  
Chief Investigator

(NAME TBC)  
(Research Nurse)
Please complete the following form by placing a tick in the box that applies and return this form in the pre-paid envelope enclosed at the address below.

I would like to attend the following focus group………

(DATE OF FOCUS GROUP HERE)...........................................

(A)........................................................................

(B)........................................................................

I am unable to attend any of the focus groups listed above but would like to be contacted about focus groups on other dates………………

(C)........................................................................

(D)........................................................................

I wish to withdraw my consent to take part in a focus group………

(D)........................................................................

Thank you

(NAME AND ADDRESS OF ADMIN SUPPORT HERE)
Appendix G

Patient interview schedule
Introduction

Good morning/afternoon. My name is (NAME OF FACILITATOR HERE) and these are my colleagues (NAME OF RESEARCH NURSE AND ADMIN SUPPORT HERE). Thank you for coming today. The idea of the focus group is to allow those present to have a relaxed discussion about something that you have experienced in the past. In this case it is the fact that you have all taken part in the Ankle Injury Tuning Fork Study, whether you presented to the (STUDY SITE HERE).

Present the purpose

We are here today to talk about your experience of receiving the tuning fork test to assess your recent ankle injury. The purpose is to give you the opportunity to discuss what you felt when the tuning fork was placed on your ankle and to get your opinion on whether you feel the tuning fork test is an acceptable method of assessing an ankle injury in the future. In addition, we would like to establish what you feel about waiting times in the emergency care setting and whether you rate receiving an x-ray or reducing the time you spend in the emergency care environment as the most important aspect of your visit.

I am not here to share information, to give you my opinions or judge what you say. Your ideas and opinions are what matter. There are no right or wrong, desirable or undesirable answers. You can disagree with each other, and you can change your mind at any time. I would like you to feel comfortable saying what you really think and how you really feel about your experience of receiving the tuning fork test. There are light refreshments available, feel free to help yourself if you have not already done so. Water is available for your use throughout the discussion.

Discuss procedure

(NAME OF RESEARCH NURSE AND ADMIN SUPPORT HERE) will be taking notes and recording the discussion so that we do not miss anything you have to say. This was explained in the patient information sheet you read when you initially attended the emergency setting with your ankle injury. As you know everything you say is confidential. No one will know who said what. No one will be identifiable from what they say or in any report written about the focus group discussion. I want this to be a
group discussion, so feel free to respond to me and to other members in the group without waiting to be called on. However, I would appreciate it if only one person did talk at a time. Please speak clearly. The discussion will last approximately one hour. There is a lot I want to discuss, so if I feel the conversation is stalling I may move us along a bit. You should have all signed a consent form today – if not please do so now.

**Participant introduction**

Now, let's start by everyone sharing their name, occupation and or favourite hobby / pastime. We’ll go round the room starting with …

**Rapport building**

In order to get to know each other a little better I want each of you to spend a couple of minutes thinking about a word that best described your personality. This word must begin with the same letter of the alphabet as your first name. Once you have thought of a word tell the person on your right the word and your first name. I.e. Feisty Fiona!

I will now ask your neighbour to introduce you using the word followed by your name. We’ll start with (NAME OF RESEARCH NNURSE HERE) and go clockwise around the room.

Okay – well done

**Interview**

Would someone like to tell me how they hurt their ankle and how long it took to get better?

Prompt: Did anyone twist their ankle walking / running? Did anyone do anything different? Can you elaborate on this? Has anyone injured their ankle before? Was it treated differently this time?

Can you tell me how you felt when you were asked to take part in the tuning fork study?

Probe: Did you need to think about it for a while or did you make up your mind fairly quickly? … … Does anyone else have a different experience … or like to add anything? … … What about family and friends – did they influence or help you with your decision?

What did the tuning fork test feel like when it was placed on your injured ankle?

Probes: did you feel a vibration … tingling … did it hurt? … … Did anyone have a different sensation / experience? … … How did this compare to when it was
applied to the uninjured ankle?

*Does anyone consider the sensation they felt to be unbearable or unacceptable?*

Probes: Can you elaborate on that ... ... Why do you think that was? ... ... What were you thinking at the time? ... ... Does anyone want to add anything?

As yet we have not analysed the data and do not know if the tuning fork test is accurate in identifying those with a fracture or not, so do not know if it will be introduced into practice or not. In your opinion do you feel that the tuning fork test should be introduced as part of the assessment of simple ankle injuries?

Prompt: Why do you feel this way? ... ... Do you feel the tuning fork test will be acceptable to other patients? ... ... Have you always felt this way or has your opinion changed? ... ... Does someone want to challenge this view or add something else? ... ... Would you advise others to have the tuning fork test?

*Would anyone like to share with the group how long they were waiting for?*

Prompt: Was this wait to see the triage nurse? ... the practitioner / Doctor to assess you? ... or waiting for x-ray? ... ... Do you consider this to be an acceptable time frame to wait? ... If not why not?

*If you had a choice what do you consider to be the most important aspect of your visit to the emergency department when attending with an ankle injury - having an x-ray ... or reducing the time you spend in the department?*

Prompt: Why do you think this is so? Have you got previous experience of this? Do you think anything could change your mind – what would this be?

*All of you present today recorded a ... ...*

......negative tuning fork test, this would mean that if the tuning fork test became normal practice for assessing ankle injuries you would not receive an x-ray. Does this alter your opinion of the usefulness of the tuning fork test?

...... positive tuning fork test, this would mean that if the tuning fork test became normal practice for assessing ankle injuries you would still receive an x-ray. Does this alter your opinion of the usefulness and acceptability of the tuning fork test?
Prompts: (for both above) – Why do you think this is so? … … Does everyone agree with this view? … … Have you always felt this way? … … Would anyone like to add anything?

Closure

Though there were many different opinions about the tuning fork test, it appears that the majority view is that it is _______. Does anyone see it differently? It seems most of you agree _______, but some think that _______.

Does anyone want to add or clarify an opinion on this? Do you all feel this is a fair summary of what has been discussed?

Is there any other information regarding your experience with or following your visit that you think would be useful for me to know?

Thank you very much for coming this afternoon. Your time is very much appreciated and your comments have been very helpful. After you have left today if there is anything that you feel you have forgotten to say please visit the tuning fork study website and leave a message on the forum. Don’t forget your travel claim form and receipts where applicable and we will endeavour to get the money reimbursed as quickly as possible. Also don’t forget to collect your £10 token before you leave today.

Thank you
Appendix H

Consent form for clinician focus group discussions
Informed consent form:

Tuning fork study focus group - practitioner participation

Thank you for agreeing to be involved in part two of the tuning fork study. This form outlines the purpose of this part of the study, describes your involvement and your rights as a participant.

The purpose of this study is:
- To gain insight into whether practitioners accept the use of the tuning fork test in clinical practice following an ankle injury, and to explore their experience of using it on patients.

The benefits of the research will be:
- To make a decision on whether it will be practicable to introduce the tuning fork test, if it is shown to be effective.
- To identify the hurdles that will need to be addressed in order to ensure it can be safely introduced into practice.
- To gain insight into practitioners views of undertaking research in the clinical setting

The method that will be used to meet this purpose is a focus group, attended by practitioners involved in the tuning fork study. A focus group allows those practitioners involved in the study to get together and discuss ideas and experiences. The focus group is being led by a facilitator who is a senior lecturer at a University. Her role is to make sure that every member of the group has a chance to share their views. The facilitator will start the conversation within the group by asking some open questions, which should then encourage a discussion to take place.

The discussion will be audio recorded to help us accurately capture the thoughts of those present in their own words. The recordings will only be heard by the members of the research team for the purpose of transcribing. Once recordings have been transcribed they will be deleted. Your direct quotes and those of others within the group will be used to write a final report. However, your name and other identifying information will be kept anonymous.

You have the right to withdraw from this part of the study at any time. In the event that you do withdraw from the study your recorded views will not be used in the final report of the study.

By signing this consent form I certify that I …………………………………..agree to the terms of this agreement. (print full name here)

Signature……………………………………………………..Date…………………………

……..
Appendix I

Interview schedule for clinician focus group discussions
Name of Moderator____

Date_______________________

Attendees_____________________________________________

Tuning Fork Study Focus Group Topic Guide - practitioners

Introduction

Good morning/afternoon. My name is (FACILITATORS NAME HERE)

I work at the University of Portsmouth. You already know (RESEARCH NURSE & ADMIN SUPPORT NAMES HERE)

Thank you for coming today. The idea of the focus group is to allow those present to have a relaxed discussion about the tuning fork study you have all been involved in.

Present the purpose

We are here today to talk about your experience of using the tuning fork test to assess ankle injuries in your clinical work place. In summary the purpose is to give you the opportunity to discuss things such as what you first thought when you heard about the tuning fork study, how you felt about approaching patients about taking part in the study, and whether you think it is acceptable as an `investigations` for the future

I am not here to share information, to give you my opinions or judge what you say. Your ideas and opinions are what matter. There are no right or wrong, desirable or undesirable answers. You can disagree with each other, and you can change your mind at any time. I would like you to feel comfortable saying what you really think and how you really feel about your experience of using the tuning fork test. There are light refreshments available, feel free to help yourself if you have not already done so, and I understand your staying for a meal afterwards!

Discuss procedure

(RESEARCH NURSE & ADMIN) will be taking notes and recording the discussion so that we do not miss anything you have to say. As you know everything you say is confidential. No one will know who said what. No one will be identifiable from what they say or in any report written about the focus group discussion. I want this to be a group discussion, so feel free to respond to me and to other members in the group without waiting to be called on. However, I would appreciate it if only one person did talk at a time. Please speak clearly. The discussion will last approximately one hour.
There is a lot I want to discuss, so if I feel the conversation is stalling I may move us along a bit. You should have all signed a consent form today – if not please do so now.

Participant introduction

Now, let's start by everyone sharing our names, how long you have been an ENP. We'll go round the room starting with...

Okay – well done

Interview

Everyone uses the Ottawa ankle rules to assess ankle injuries don’t they?

- Prompt: If not why not?...Do you know the types of ankle injuries they are used on?...Has your opinion on this changed since you have taken part in the tuning fork study?...Have you looked at research to do with the Ottawa ankle rules?...Do you feel confident in them?...Do you trust them?...Do you think they are accurate enough?...Are you happy to turn patients away without x-ray if they are assessed as Ottawa negative?...Do you feel your assessment using the Ottawa ankle rules changed during the course of the study?

Would someone like to start by telling me what they first thought when Anne came to you and told you about the tuning fork study?

- Prompt: What another study!...What about time?...We are far too busy!...Does anyone else have a different view?...Does anyone want to add anything?...Did anyone feel this is a waste of time?...Did anyone feel they did not want to take part in this?...Have you been involved in clinical studies before?...Did this one make you feel different?...If so, How?...Were you excited?

Do you feel you had sufficient training in the use of the tuning fork test before the study started?

- Probe: Who did the training?...Were you offered additional training if you needed it?...Were the research team available for questions?...Were you given anything to read about the reasons for doing the study?...Could the training have been done in a different way?...What else could have been done to make the training better?

How did you feel about approaching patients about the study?
• Probes: Were you anxious?.....Did your feelings change the more patients you approached?.....Did you ask every patient who met the criteria or did you ‘select’ certain ‘types’?....Why did you select some patients and not others….Did this depend on workload?.....Patient attitude?...Time of day / shift?.....Does anyone want to add anything?

Did you feel confident doing the tuning fork test? Did the research team give you sufficient support throughout the study?

• Probes: If not, why not?..... What would have helped your confidence?......What further support would you have liked?

As yet we have not analysed the data and do not know if the tuning fork test is accurate in identifying those with a fracture or not. However, in your experience do you think it is accurate?

• Prompt: Did you have any false negatives (tuning fork negative, x-ray positive)? .....If so, did the x-ray change your management of the injury?.....

What do you think the patients thought about the tuning fork test?

• Prompt: Did any express concerns or reservations?......Did anyone refuse to take part?.... Did they give any reason?.....

Do you feel the tuning fork test will be acceptable to other patients in the future?

• Prompt: Have you always felt this way or has your opinion changed? … Does someone want to challenge this view or add something else? … ….Would you advise other clinicians to use the tuning fork test?

The Ottawa ankle rules when used alone are about 95-100% sensitive at ruling out a fracture but only about 15% specific at ruling in a fracture. What this means is that the Ottawa ankle rules are accurate at identifying patients that require an x-ray but of those x-rayed a fracture is present in only 15%. Bearing in mind that very few tests are 100% accurate what do you feel would be a high enough specificity for the tuning fork test to be acceptable to clinicians?

• Prompt: Do you understand sensitivity and specificity?......Currently assessing ankle injuries using the Ottawa ankle rules means that 15 out of every 100 are identified as having a fracture – do you feel this is sufficient……Between 2-5
people out of 100 assessed as Ottawa negative are later found to have a fracture – is this satisfactory?.....Would you accept the tuning fork if they missed any clinically significant fractures?

The data has not yet been analysed. Remember that if the tuning fork is shown to be effective then patients recording a ‘negative’ tuning fork test will not get an x-ray in the future. In your opinion do you feel that it should be introduced as part of the assessment of simple ankle injuries within an ED / MIU setting?

- Prompt: Have you always felt this way or has your opinion changed? … Does someone want to challenge this view or add something else? …Do you feel it would be more beneficial in GP surgeries etc?

What are the advantages and disadvantages for practitioners, the NHS and patients?

- Prompts: What about complaints, legislation, etc?.....Finances?......Time?......Does anyone else have a different perspective? … …Would anyone like to add anything?

Closure

Though there were many different opinions about the tuning fork test, it appears that the majority view is that it is _______. Does anyone see it differently? It seems most of you agree ______., but some think that ____.

Does anyone want to add or clarify an opinion on this? Do you all feel this is a fair summary of what has been discussed?

Is there any other information regarding your experience that you think would be useful for me to know?

Thank you very much for coming this afternoon. Your time is very much appreciated and your comments have been very helpful. After you have left today if there is anything that you feel you have forgotten to say please visit the tuning fork study website and leave a message on the forum.

Thank you
Appendix J

Sensitivity analysis
As qualitative analysis is an interpretative process, the preconceptions, assumptions and ‘worldview’ of the researcher are likely to influence the process and any emerging theory, despite use of rigorous approaches. A reflexive account is an honest attempt by the researcher to declare their conceptual journey through the research (Lacey & Luff 2009). This sensitivity analysis attempts to identify how my preconceived ideas and thoughts may have influenced the qualitative data analysis.

I am a female aged 46 years and am currently employed full-time as a nurse consultant. I qualified as a registered nurse in 1986 and quickly realised ED was where I wanted to focus my career. After a short time working in theatres, an acute medical ward, and an orthopaedic ward I secured a staff nurse position in the ED. At this point in my career ENPs were new and slowly being introduced across the NHS but as early as 1987 I knew I wanted to follow this career pathway.

After a career break between 1995 and the year 2000 I returned to the nursing profession, again within an ED, and very quickly undertook my registered children’s nursing. I was then lucky enough to secure a trainee ENP role. It was during my ENP training in 2003 that I first came across the OARs. I was informed about when and how to apply them and the meaning of each criterion. I continue to use them today on ankle injuries, know them without reference and could recite them in full prior to the start of the study. In my opinion the OARs help reduce the number of x-rays requested but on occasions I have known patients assessed as OAR negative return only to find they have a fracture. I therefore know they are not 100% accurate. I have also had experience of where a patient has been OAR negative but clinical instinct has told me to x-ray them only to find they have a fracture. These patients often have little or no swelling but tend to complain of pain over a small area, compared to ankle sprains where the tenderness tends to be more widespread. Prior to undertaking this study I was unaware of other ankle rules and am interested in pursuing the Buffalo rules as an alternative to the OARs.

When I first heard about the tuning fork test in the Dissmann and Han paper I was intrigued. I felt this was a novel way to try and reduce ankle x-rays whilst continuing to use the widely validated OARs. I discussed the use of the tuning fork test with colleagues and was surprised to hear that it is routinely used in the armed forces for stress fractures. Shortly after publication of the Dissmann and Han paper I received a referral from a military medical officer for an x-ray of an ankle based on a positive tuning fork test to the patient’s ankle.

Undertaking the systematic review into the use of the tuning fork test I was surprised to see how long ago it was that this was first discussed in the literature. I assumed that the invention of x-rays had meant diagnosis with a tuning fork was condemned to history. However, I strongly felt that there was a place for the tuning fork in today’s society. More and more emphasis is being placed on the importance of patients receiving the right care, in the right place, by the right person and attendances to EDs have risen significantly in the past 10 years. Implementing something like the tuning fork into GP practices may reduce the need for patients to come to the ED. Experience has taught me that...
patients don’t want to come to ED unnecessarily but many are confused where they can go with a minor injury. Patients do not understand the difference between a minor injury unit, a walk-in-centre or an urgent-care-centre. What they do know is that very few people get turned away from an ED, but there is increasing pressure on the staff within EDs to do this. I did not attend the focus group discussions as I felt I could be accused of biasing the discussion but I was pleasantly surprised when I listened to the recordings to hear patients talk about the need to introduce the tuning fork into GP surgeries. This made me realise that preconceived ideas that clinicians sometimes have about patients are not always true.

There is the perception amongst clinicians that patients expect an x-ray and I myself have witnessed this, but I also feel this is fuelled by clinicians who are not adequately trained in the assessment of injuries raising patient expectations. I have experience of patients being referred by other clinicians to the ED for an x-ray when clearly they have a soft tissue injury. Educating these patients has always been a passion of mine, but sometimes work pressures dictate that I don’t have the time to spend discussing the injury with the patients and it can sometimes be quicker to send the patient to x-ray.

I am naturally inquisitive and enjoy challenging practice and as such thrive on the work of a consultant nurse which challenges the boundaries of the nurse-doctor profession. Since I started working in the ED I have noticed how the doctors have become more and more deskillled in the management of minor injuries. This is more evident since the introduction of new junior doctor training. However, some of the responsibility of this lies with the advent of the ENP role. ENPs are now employed in the majority of EDs to manage minor injury and as such doctors are becoming deskillled in this area.
Chapter 11 - COPIES OF PUBLICATIONS

1 – A4 Word copy of poster submitted to the International Conference for Orthopaedic Nurses in Dublin, 2010

2 – Powerpoint presentation at the College of Emergency Medicine & the RCN Emergency Care Association conference 2011

3 – A4 word copy of the poster submitted to the RCN Emergency Care Association conference 2011
The Tuning Fork test- An accurate and efficient method of improving the diagnostic accuracy of the Ottawa Ankle Rules.

A Welling¹, S Cooke¹, M Archer¹, A Dewey², B Higgins², S Carss¹
¹Portsmouth Hospitals NHS Trust, United Kingdom

INTRODUCTION & AIMS.
*The Ottawa ankle rules (OARs) are routinely used to assess simple twisting ankle injuries in emergency care settings and determine the need for x-ray.

*However, the incidence of fractures in those x-rayed is only 15%(1)

*The primary aim of this study is to assess whether a tuning fork can increase the diagnostic accuracy of the OARs.

*Secondary aims include identifying patients views on the use of the tuning fork, whether the tuning fork could reduce the number of ankle x-rays requested, and reduce the time a patients spends in the emergency care setting.

METHODOLOGY.
*A Multi-site mixed methods study comprising a quantitative diagnostic test and qualitative focus groups

*The study is taking place in three Emergency Departments and One Minor Injuries Unit in the UK.

*The diagnostic test will assess the accuracy of the OARs when used in conjunction with the Tuning Fork Test and this will be compared to routine AP and Lateral ankle X-rays

*The focus groups will assess patient and practitioner views on the acceptability of the tuning fork test in clinical practice.

SAMPLE.
*1300 participants are needed
*Simple Eversion/ Inversion injuries only
*Age 12yrs and Over
*Ottawa Ankle Rules Positive

INTERIM RESULTS
*Study commenced July 2009.
*To end August 2010 total 795 participants recruited
*No analysis of diagnostic study yet performed
*Analysing a small sample of 250 participants identified as tuning fork negative and x-ray negative the average wait from being sent to x-ray to being discharged was 29 minutes.

SPONSOR / FUNDING / ETHICS
*Sponsored by Portsmouth Hospitals NHS Trust
*Funded by the NIHR RfPB Ref: PB-PG-1207-15022
*UKCRN Portfolio database ISRCTN18630663
*Study approved by Southampton & SW Hants REC ref: 09/H0502/57
*Study website: www.tuningfork.org.uk

REFERENCE
• Bachman, Kolb, Toller, Steurer, Rist. (2003). Accuracy of the Ottawa ankle rules to exclude fractures of the ankle and mid-foot: systematic review. BMJ [electronic version], 326 (7386), 417-
# Tuning fork testing on ankle injuries: does it improve the accuracy of the Ottawa ankle rules?

Anne Welling  
Consultant nurse in emergency care  
5th August 2011

## Background & rationale for study

- Ankle injuries account 8% of all minor injuries  
- 1.5 million UK ED attendances annually for sprains  
- Ottawa ankle rules designed to identify clinically significant #s  
- OARs = Good sensitivity, poor specificity  
- Fracture prevalence 15%  
- X-ray adds 30 mins to LOS  
- Can accuracy be improved?

## Methodology

- Mixed methods  
- Screening test – OARs & tuning fork  
- Focus groups – explore patient & clinician views  
- 4 study sites  
- June 2009 – June 2011  
- 1300 target recruitment  
- Funding from CCF NIHR RfPB

## Sample

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Inclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>12+</td>
</tr>
<tr>
<td>Gender</td>
<td>OARs+ malleolar zone</td>
</tr>
<tr>
<td></td>
<td>Simple twisting MOI</td>
</tr>
<tr>
<td></td>
<td>No obvious deformity</td>
</tr>
<tr>
<td></td>
<td>Able to give informed consent</td>
</tr>
<tr>
<td></td>
<td>No Hx peripheral neuropathy</td>
</tr>
</tbody>
</table>

## Data collected in addition to TFT and x-ray result

<table>
<thead>
<tr>
<th>Data collected</th>
<th>Inclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Randomisation of testing</td>
</tr>
<tr>
<td>Gender</td>
<td>Study site</td>
</tr>
<tr>
<td>Injured ankle</td>
<td>Clinician performing test</td>
</tr>
<tr>
<td>Degree of swelling</td>
<td>TOA</td>
</tr>
<tr>
<td>Bony prominence visible</td>
<td>Time to assessment</td>
</tr>
<tr>
<td>Distracting injuries</td>
<td>Time to discharge</td>
</tr>
<tr>
<td>Previous injury</td>
<td></td>
</tr>
</tbody>
</table>

## Results

- 1358 recruited – 45 removed  
- 1313 analysed  
- Mean age: 34 years, range: 12-91  
- 730 male / 583 female  
- Ethnicity: 98% white  
- Right ankle: 51%  
- 210 ‘d = 16% (39 deemed not clinically significant)  
- LM & DFS = most common #s  
- 113 clinicians (60 nurses, 53 doctors)
Results (continued)

<table>
<thead>
<tr>
<th>TFT+</th>
<th>X-ray+</th>
<th>X-ray−</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>145</td>
<td></td>
<td></td>
<td>172</td>
</tr>
<tr>
<td>27</td>
<td></td>
<td></td>
<td>1141</td>
</tr>
<tr>
<td>702</td>
<td>557</td>
<td>584</td>
<td>1313</td>
</tr>
</tbody>
</table>

Correct 729
Accuracy 56% (CI 53-58)
Sensitivity 84 (78-89)
Specificity 51 (48-54)
NPV 96 (94-97)
X-ray reduction 47%

Missed # include
7 managed as STI
10 DFS (7 tender LM)
6 LM/talus avulsions 3-9mm
1 LM avulsion17mm
1 LM 25 days post injury
2 bi-malleolar

Results (continued)

Predictors of a correct TFT after adjustments for confounding

<table>
<thead>
<tr>
<th>P value</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.020</td>
</tr>
<tr>
<td>Clinician role</td>
<td>0.003</td>
</tr>
<tr>
<td>Gender</td>
<td>0.078</td>
</tr>
<tr>
<td>Male gender</td>
<td>1.247</td>
</tr>
</tbody>
</table>

Correct test more likely with increasing age, male gender and ENP performing test

Results (continued)

*Mean wait to be seen – 49 minutes (0-240)
*Mean time from assessment to discharge 62 minutes
*63% discharged within 2 hours
*99% discharged within 4 hours

Patient focus groups
*TFT acceptable with information
*rated reduced waiting times more important than x-ray
*Recommend use by GP, pre-hospital, remote areas

Results (continued)

Clinician focus group
*Sceptical at first
*Recruitment dependent on workload
*Patients want / demand x-ray
*Doctors likely to x-ray more
*Easy to learn & use
*Must compare with ‘good’ ankle
*Would not be accurate to be used alone

Summary

*Multiple sites & clinicians
*reduces x-rays by 47%
*Improves specialty of OARs
*NPV 96%
*Age, gender & role of clinician predictors of correct & + result
*Swelling, distracting or previous injury not predictive of correct result
*Accepted by patients & clinicians
*easy to learn & use

Limitations & recommendations

*Only used on simple twisting mechanisms
*’Malleolar zone’ only – excluded foot
*12 years and over
*’White’ ethnic origin
*Further studies in ethnic groups, children, and foot injuries recommended
Tuning fork testing on ankle injuries: Does it improve the accuracy of the Ottawa ankle rules?

A' Welling\textsuperscript{1}; A Dewey\textsuperscript{2}; B Higgin\textsuperscript{2}; S Cooke\textsuperscript{1}; M Archer\textsuperscript{1}

\textsuperscript{1}Portsmouth Hospitals NHS Trust and \textsuperscript{2}University of Portsmouth

Background
* Ankle injuries account for 8\% of all minor injuries
* 1.5 million UK ED attendances annually for sprains
* Ottawa ankle rules designed to identify clinically significant \#s
* OARs = good sensitivity (95\%), poor sensitivity (30\%)
* Fracture prevalence in those x-rayed = 15\%
* X-ray adds 30 mins to time in ED
* Can accuracy of OARs be improved?

Methodology
* Mixed methods
* Screening test – OARs & tuning fork – interpreted blind
* Focus groups – to explore patient & clinician views

Results
* 1358 recruited – 45 removed – 1313 analysed
* 113 clinicians involved in recruitment

<table>
<thead>
<tr>
<th>Patient demographics (number unless stated)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (range)</td>
</tr>
<tr>
<td>Male gender</td>
</tr>
<tr>
<td>‘White’ ethnicity</td>
</tr>
<tr>
<td>Right ankle injured</td>
</tr>
<tr>
<td>Bony prominence visible</td>
</tr>
<tr>
<td>Distracting injuries</td>
</tr>
<tr>
<td>+++/++++ swelling</td>
</tr>
<tr>
<td>Previous injury to ankle</td>
</tr>
<tr>
<td>Patients with clinically significant #</td>
</tr>
<tr>
<td>Mean time from arrival to assessment</td>
</tr>
<tr>
<td>Mean time from assessment to discharge</td>
</tr>
</tbody>
</table>

Relationship between tuning fork test and x-ray diagnosis

<table>
<thead>
<tr>
<th>Tuning fork</th>
<th>Xray +</th>
<th>Xray –</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuning fork +</td>
<td>145</td>
<td>557</td>
<td>702</td>
</tr>
<tr>
<td>Tuning fork –</td>
<td>27</td>
<td>584</td>
<td>611</td>
</tr>
<tr>
<td>Total</td>
<td>172</td>
<td>1141</td>
<td>1313</td>
</tr>
</tbody>
</table>

Accuracy | 56\% (95\% CI 53-58)
Sensitivity | 84\% (78-89)
Specificity | 51\% (48-54)
Negative predictive value | 96\% (94-97)
Potential x-ray reduction | 47\%
Potential reduction in length of stay for tuning fork negative patients | 66 mins

Predictors of a correct tuning fork test after adjustments for confounding using a multiple logistical regression model:

<table>
<thead>
<tr>
<th>P value</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Clinician role</td>
<td>0.003</td>
</tr>
<tr>
<td>Gender</td>
<td>0.078</td>
</tr>
</tbody>
</table>

| Age | 1.020 |
| Clinician role | 1.595 |
| Male gender | 1.247 |

Conclusion
The tuning fork test can be used in conjunction with the Ottawa ankle rules to assess simple twisting ankle injuries to determine the need for x-ray, resulting in a reduction in x-rays and a reduction in length of stay. However, this is at the expense of missing some clinically significant fractures. The test is accepted by patients and is easy to learn and apply in the clinical setting.

Recommendations
Further research into the use of the tuning fork test in different ethnic populations, younger age group and the ‘foot zone’ of the OARs