Draft Protocol – Pre-ethics submission

Synthesising a clinical Prognostic Rule for Ankle INjuries in the Emergency Department (SPRAINED)

1. Planned Investigation

Introduction

Incidence and costs

Ankle sprains are one of the most common musculoskeletal injuries. Up to 1.5 million people a year in the UK attend ED with a sprained ankle. The vast majority of sprains are of the lateral (outside) ligaments, and vary from minor stretching (Grade I, mild sprain) to a complete tear (Grade III, complete sprain). Recovery can be protracted particularly for more severe injury. The health costs of managing moderate to severe sprains with a tubi-grip support are on average £135 per sprain, although costs of operative treatment and/or physiotherapy are considerably greater. Including work absence the costs are approximately £900 per patient (2005 prices), with an average of 7 days work absence.[1]

Recent systematic reviews conclude that approximately 30% of people have problems one year after an ankle sprain depending on the outcome measured and perhaps more importantly, the sampling frame.[2,3] One sequel, chronic ankle instability (CAI) is implicated in the development of ankle osteoarthritis, even without an acute osteochondral lesion.[4] Many studies are restrictive in their sampling frame, either to elite athletes and exclude younger and older people. Studies also have variable inception and follow up points which further complicates interpretation. We report an estimate of 30% for poor outcome at 9 months in the Collaborative Ankle Support Trial (CAST) dataset.[1]

Studies agree that recovery plateaus around 9 months, and that residual disability after this point is likely to be persistent.[5] Persisting severe symptoms are generally reflective of incomplete repair of the ligament/soft tissue structures, muscle weakness or injury not identified at initial consultation.

Usual clinical pathway

Assessment of the injury in the hyper acute phase is challenging, the ankle is often so swollen and painful that it cannot easily be touched. Most patients are advised to rest, elevate, apply ice and compression, and are often issued with crutches. Use of X-ray has been effectively controlled through the Ottawa guidance,[6] with only the most concerning injuries being X-rayed (or equivalent) to exclude fracture. Where clinicians are concerned about the degree of injury, most providers operate a system of review within one week in a trauma or equivalent injury service. This time frame allows dissolution of some swelling, and greater certainty in ascertainment of injury severity and presence of other significant mechanical derangement.[7] Treatment options at this stage include further watchful waiting, diagnostics, intensive physiotherapy or immobilisation. Surgery may be considered at this stage, although most centres would initiate a test of conservative treatment first. We have previously published a survey of practice, which whilst quite an old publication remains an accurate reflection of the current management.[8]

Value of a prognostic tool

A prognostic tool is indicated to identify patients likely to experience poor outcome. There are effective treatments available, and it is generally accepted that early intervention is preferable. There are clearly several steps where better prognostic information could yield benefit to the NHS and to patients as follows; [1] Increase the certainty that an early review is merited , and avoid unnecessary appointments [2] Better target treatments and diagnostics, earlier in the recovery pathway [3] Be re-assured that patients not followed up are on a positive recovery trajectory. Volume is a key issue for sprains. Cost savings will accrue if treatments are more efficiently targeted.

Any prognostic tool needs to be simple to complete in the ED. Ideally a tool which requires a once only administration is preferable, although there are several possibilities which we can explore in the CAST data set. The issues that we will explore during development will be variable selection, timing and method (self-report versus clinical examination).

Requirements of a prognostic tool
To be considered useful, a prediction tool should be clinically meaningful, accurate (well calibrated with good discrimination) and generalizable (have been externally validated). Many prognostic tools are developed using datasets that are too small, are not sufficiently generalizable, have questionable methodological quality (in particular no internal or external validation) and use inadequate statistical methods (e.g R^2 from multivariate regression).

Aims and objectives

The aim is to develop and validate a prognostic tool for use in EDs for patients with acute ankle sprain in order to identify those in whom recovery may be substantially prolonged or incomplete and additional investigation/treatment is indicated. Specifically we will;
Stage 1: Develop and internally validate a prototype prognostic tool to detect risk of poor outcome in people who presented with moderate and severe ankle sprains in the CAST data set. We will explore the trade-offs in prognostic ability of different approaches.
Stage 2: Undertake a systematic review, or review of reviews, to ascertain variables not included in the CAST data set which might add additional prognostic information and to gain clinical and PPI input into the final selection of models and variables for testing, and the format of the tool.
Stage 3 Externally validate and optimise the tool in a new cohort of 675 patients recruited from a representative group of NHS emergency departments and minor injury units.

Existing Research

We have undertaken scoping searches in the preparation of this application to minimise the possibility of research waste, inform us of the evidence base for treatment options for ankle sprains and provide a scope on the size of the literature on predictors.

Existing prognostic tools
There are no validated prognostic tools for acute ankle sprain and most of the existing studies are of limited use due to highly selective patient populations (exclusion of some the more severe type of injury, older patients and/or sole inclusion of athletic/military populations). There are no studies on various international research registers (Current controlled trials, ClinicalTrials.gov, clinicaltrialsregister.eu).

Hiller et al[9] provide a good systematic review of factors associated with the risk of sustaining an ankle sprain but there are very few studies evaluating risk of poor recovery after the injury. Other than recurrent sprain, few studies of post-injury recovery have considered these pre-dispositional factors. In 2008, Van Rijn et al[2] published a systematic review of the clinical course and prognostic factors for recovery following ankle sprain. They found just one eligible study [10] which concluded that high levels of sports activity was a prognostic factor for residual symptoms (n=150).

We have re-run the searches used by Van Rijn (up to December 2013) and located an additional 6 studies investigating prognostic studies for outcome following ankle sprain, although these are not validated prognostic tools. De Bie et al[11] evaluated clinical and self-report measures at 2 and 4 weeks in 35 ED patients. A combination of function score, palpation and initial clinician rating of severity were predictive of a clinician judgement of "healed" at 4 weeks. Wilson and Gansneder[12] evaluated impairment and self-reported activity limitation measures for disability duration in 21 athletes. Self-report activity limitation measures were the strongest predictors of time to return to sport participation. Van Middelkoop[13] followed 102 patients in ED and primary care up to 3 and 12 months. They considered age, sex, BMI, treatment, setting (GP or ED), injury grade (mod/sev), swelling, Ankle Function Score, Work load, sport load, and pain during walking but were unable to identify any predictive factors. There was a suggestion that re-injury at 3 months was predictive of 12 month recovery score, but this analysis was performed in a subset of participants. Gerber et al[14] performed a study evaluating young personnel at a US military academy and found that signs of syndesmosis sprain were most predictive of poor outcome at 6 months. Most recently, O'Connor et al[15] investigated predictors in a UK population, recruited from ED/ sports injury clinic, following 85 patients with mild or moderate acute ankle sprains at 4 weeks and 4 months. They found that medial joint line palpation, pain during weight bearing ankle dorsi-flexion at 4 weeks explained about 50% of variance in outcome at 4 months. There were significant limitations to this study; inclusion of only mild and moderate ankle sprains, an upper age limit of 65, and a substantially underpowered analysis. Finally, data reported by Akacha and Hutton,[16] based on a re-analysis of the CAST data set demonstrated the very substantial effect that age has on recovery.

Clinical opinion and guidance

Practice recommendations have been published by European groups. In Holland, Kerkhoffs et al[17] developed a practice recommendation but this is intensive and goes well beyond the evidence. Polzer et al[3] have developed a prognostic algorithm and treatment pathway, but again substantial sections are expert based judgements of a poor evidence base. It is widely recognised that a properly developed, well validated prognostic tool could help better target treatment and improve outcomes for patients.

Evidence for treatment effectiveness

There are treatment options available for people who have poor prognosis. The most solidly evidenced based is physiotherapy.[18] Others options include surgical reconstruction of ligaments.[19]

Stage 1 – Developing a multivariable prognostic model from an existing dataset for assessment of risk of poor outcome

We will identify variables for the prognostic tool using multivariable logistic regression modelling of the CAST data set. CAST is the largest trial of interventions for moderate to severe acute ankle sprains to date (worldwide), by a considerable margin (n=584).[20] It was a four arm, individually randomised trial, comparing three types of mechanical support in comparison to standard care (tubigrip) in eight emergency departments. All people who attended ED with a moderate or severe sprain (Grade II and III) were approached, and we randomised all those willing to participate. We used a pragmatic indicator of injury severity (weight-bearing three days after initial attendance). We have a minimum anonymised data set on those declining the trial and those with mild injuries, but no follow up on these individuals. For participants of the trial, we have data at ED presentation, 3 to 5 days later (when injury severity can be more accurately ascertained and we randomised), then at 1, 3 and 9 months (follow up).

We collected initial data on 1522 patients with all grades of sprains, and have follow up on between 584 and 441 moderate and severe sprains, depending on the point of follow up. There is data on a

comprehensive number of predictors, including those identified as potentially important by clinical guidelines/consensus, and in multivariable analyses (See Appendix 1).

Concerns with using a randomised sample for a prognostic algorithm would be around sampling (transportability), and the random assignment of treatment.[21] We contend that we had a good selection of EDs and generated a sample with good representation of ethnic minorities, gender, age range, physical activity and sporting ability. The great majority of trials and observational studies conducted to date, even those in the UK,[22] have excluded people of older age (usually set at 60 years). In the CAST data we found age to be a strong predictor of outcome. Given that it is unethical to withhold treatment, any prognostic modelling process is going to have to consider treatments provided (regardless of whether randomly allocated). The treatments in CAST were tubigrip, a removable air-cast brace, a removable Bledsoe boot and a 10 day below knee cast. All of these treatments are in use in the NHS but prior to CAST had not been evaluated in an RCT. The 10-day below knee cast was moderately effective in accelerating recovery in the early phases of treatment (measured up to 3 months), but made no difference to 9 month outcomes (30% of people had residual problems in each arm). We did not organise any other treatments or interfere with the normal clinical pathways in CAST, although we have collected information on health services resource use throughout the follow up period. We see it as a strength that we are aware of the treatments received. The prognostic modelling will be focused on the 9 month outcome, where treatments did not affect outcome. We may use data from the 1 month follow up to determine the importance of delayed data in the prediction rule. Moreover, we envisage the predictive effect of interventions to be small compared to the important predictors such as age, sex and other key predictors.[21] We will undertake sensitivity checking by modelling across and within treatment arms. Treatment will also be included as a candidate predictor during model development (but may be subsequently omitted during the final multivariable modelling). If there is some sensitivity to treatment allocation, we will combine the treatment arms where no treatment effect was observed and base our main conclusion on this sub-set of the data.

The main limitation of the CAST data set is that we excluded people who had minor sprains and were able to weight bear by day 3 to 5. We have some informative data (see above) and do not anticipate this to be a great problem. Grade I sprains are minor and have a good prognosis.[3] The final point to note is that the 9 month outcome in CAST is ascertained directly from the patient, and not clinical examination.

Candidate predictor variables include age, sex, severity of sprain at initial attendance, initial pain intensity, history of previous injury, anterior draw test, talar tilt test, ability to weight bear, amount of weight bearing, and severity of presenting symptoms. We will model the impact of a review at about 5 days when tests can be easier to complete and have some evidence for increasing accuracy of diagnosis. We have already published work demonstrating that recovery is influenced by age and gender.[16]

There is currently no consensus on the best approach to developing a prediction model, however, we will use a transparent process that implements appropriate statistical methods and adheres to current methodological recommendations.[23,24] We will choose a priori up to 15 candidate predictor variables for inclusion in a multivariable logistic regression model (recovered or not). Simulation studies examining predictor variables for inclusion in logistic regression models suggest at least 10 events per candidate predictor to avoid over fitting,[25] whilst others have suggested this figure could be as low as 5.[26] A backwards selection procedure will be used to select which of the candidate predictor variables should be included in the final prediction model (with $p < 0.2$ conservatively taken to warrant inclusion and prevent over fitting). All continuous predictors will be kept as continuous in the modelling to avoid any loss in power (e.g. by dichotomising).[27] Any continuous

predictors exhibiting a nonlinear relationship with the outcome will be considered for modelling using fractional polynomials.[28] There is missing data in the CAST data set, although the baseline predictor set is near complete, and final follow up is greater than 76%. To avoid excluding patients and thereby reducing the sample size, multiple imputation will be used to impute missing values, under a missing at random assumption.

The number of available predictors plausibly related to the outcome could be as many as 30. We will reduce the number of predictors for inclusion into the multivariable modelling by examining whether predictors can be combined (due to multicollinearity), based on either subject knowledge, statistical clustering techniques, complexity measurement and completeness of measurement.[23,29] Predictors will also be considered for omission if the distribution of the predictor is narrow, thereby unlikely to contain sufficient predictive information and face-validity.

We will run and compare a series of models to compare the inclusion of objective clinical examination variables supplementary to self-report items, as well as the added value of additional information at different time points. We will explore the utility of the 4 week data from the CAST dataset to replicate findings of previous prognostic models generated from the baseline and follow up data of clinical trials.[13,15] We will also look to see whether a targeted approach at 4 weeks is helpful (for example picking up those who self-report persisting problems).

A prediction model that is based on self-report variables is particularly attractive as it allows for the possibility that patients can self-complete the tool without the need for clinical consultation. PPI feedback has emphasised the inconvenience of trying to get back to hospital when you have an acute injury (at 3 to 5 days after initial presentation), although it has also confirmed that if the information was sufficiently valuable then people would be willing to consider this.

Outcome that we will predict:
The key factors that signify a poor outcome after ankle sprain are instability of the joint, which is typified by recurrent sprains and or a significant lack of confidence in the ankle (a persistent feeling of giving way), with or without chronic pain. There is good consensus in international guidance that any of these outcomes would merit further investigation and intervention.
For the development dataset we will define poor outcome at 9 months as either

- Severe or persistent chronic pain/functional difficulty,
- Significant lack of confidence in the ankle
- Recurrent sprain

These items were collected by self-report in the CAST study. The selection of these variables as outcome indicators is supported by evidence from van Rijn[30] who reported recovery was most closely associated with improvements in pain and giving way. Wikstrom et al[4] report pain and instability are of greatest concern to patients. We have examined the CAST data set and the event rate is 30%. Given the event rate and number of predictors we will carry forward, we have more than adequate numbers for modelling.

Internal validation, score and format derivation
We will internally validate the models using bootstrapping, and adjust for over fitting. We will simplify the presentation of various models to a scoring system.[23,31] This can be more challenging than initially appears, but we will enlist the help of our PPI reps and clinical colleagues in developing a rational tool for clinical/patient self-completion.

The prognostic models will produce a risk score (probability) for each patient. The statistical measures of calibration or discrimination do not capture the clinical usefulness of the prediction model (i.e. how does using the prediction model improve on a default policy of no prediction

model). This will be examined using decision curve analysis, which calculates the net benefit of using the model compared to not using the model,[32] evaluating and determining ranges of cut-off values for decision-making.

Ethics

The CAST study had full ethical and other approvals, and the consent taken covers the intended use of the data.

Stage 2 – Evidence synthesis from literature and expert consensus process to refine prognostic tool

The CAST data set whilst comprehensive is not exhaustive. For example, we included some physical examination tests, but there are alternative tests available and since CAST was completed, technologies around gait assessment have moved on considerably. Scoping searches completed (see section Existing Research) do not suggest major deviation from the variables collected in CAST, however we cannot (and should not) rule out the need to collect additional variables.

2a – systematic review

We will conduct a systematic review to identify risk factors for poor outcome following acute ankle sprain to identify any additional variables that should be considered in the external validation study. We will perform a search of international online databases [MEDLINE, CINAHL, PsycINFO, Embase] from inception to date. There will be search strings for condition, body area and methodology (Table 1). The search strategy used by van Rijn et al[2] will be modified to maximise sensitivity.[33] We will investigate bibliographies of retrieved studies and consult with experts in the field. No language restrictions will be applied in the searches. Two reviewers will independently screen titles or abstracts, screen retrieved studies for inclusion, extract data and assess study quality.[34] If agreement is not achieved a third reviewer will adjudicate. A standardised data extraction form will be developed and utilised, and the QUIPS checklist for evaluating prognostic tools.[35,36] If appropriate we will compute pooled estimates for risk factors. We will use a similar approach to one we developed for a review of risk factors in whiplash injuries that enable us to classify the level of risk and certainty associated with various risk factors.[37] Those factors for which risk is moderate or high, and have a reasonable degree of certainty (i.e. have been replicated in more than one cohort), we will suggest as additional candidate variables for the final tool (see later methods)). The systematic review will be registered on the PROSPERO international database of prospectively registered systematic reviews in health and social care (http://www.crd.york.ac.uk/PROSPERO/).

Table 1 - Search terms for systematic review

Search String	
Condition	Injur* OR sprain OR strain OR inversion
Body area	[Ankle OR talocrural OR talofibular OR calcaneofibular] AND ligament
Methodology	Incidence (MeSH) OR explode cohort studies (MeSH) OR prognos* OR prospect* OR predict* OR course

2b – Expert consensus process

At this stage of the project we are likely to have a series of candidate tools, options on the presentation of tools, and additional variables/tests to consider for inclusion. The decisions that will need to be made will be complex and require input from clinicians and patient representatives. For example, we may find a small improvement in prognostic accuracy with addition of measures at about 5 days, but the implication is inconvenient re-attendance at hospital for patients.

There are several methods that we could use to gain consensus or information on preferences (including Delphi methods, Discrete Choice Experiments, and face to face methods).[38] We have opted for a method (modified nominal group technique (MNGT)) which is well recognised, can be operationalised with relative ease, has a solid audit trail, and is based on face to face consensus building. We have previously used this technique to build consensus on injury outcome measures.

The MNGT is centred around three phases. First is the preparation of briefing papers. For these we will use the systematic review detailed in the preceding section (or short report of), a lay report of the modelling elements completed in Stage 1, and a number of alternative template presentations for the prognostic tool. Additional predictive factors where there is consistent evidence of moderate or high risk will be considered by the participants of the MNG to decide if the factors are sufficiently distinctive or potentially simpler alternatives to existing elements of the prognostic tool derived from the development dataset. The MNG will evaluate and advise on the practicalities of extra measures. The consensus element is achieved by small groups, with independent facilitation, and a pre-specified set of questions. Consensus is achieved in two stages. The first round is identification of issues, and general discussion. The second round is resolution and consensus. We will use experienced and independent facilitators and ensure all opinions are heard and considered. There is no formal estimation of the degree of consensus and we will not set a formal rule in relation to this. The face to face discussions will be invaluable in decision making.

The participants will be a maximum of 12 clinicians (including ED doctors, trauma/sports physicians, physiotherapists, ED nurses, and minor injury unit staff), and 12 patients (representing a range of age, gender and physical activity/sport participation). The tool will then undergo final revision and common sense testing before release for external validation. Our plan will be to present two schedules of data collection for completion (excluding data on sample characteristics that are not contained in either schedule). The first will be prognostic tool as generated by the development data set. These will require a degree of formatting to convert data collected in a clinical research form, to an easy to complete prognostic tool. The additional predictive factors will also be formatted in a similar and completely complementary manner, but collected on the separate schedule (and after completion of the primary tool). The two schedules will be collected by the same clinician. In this way if any of the additional factors is confirmed as important, we will add them into the final tool, causing least disruption to the process of validation. If there is more than one additional predictive factor, we will randomise the order of data presentation across participants to gain insight into whether the order or opportunity to provide additional data in anyway biases the sample.

All participants of the consensus exercise will provide consent to participate. PPI members will be paid for their time and contribution, clinicians will not.

Stage 3: Externally validate and optimise the tool in a new cohort of patients recruited from NHS Emergency Departments

We will externally validate the prognostic tool developed in stages 1 and 2 in a new prospective cohort. We will recruit 675 people from 5 EDs, from a pool of 7,500 to 10,000 ankle sprains per annum at a conservative rate of 15 sprains per centre per month. We have provisional agreement from an additional 5 EDs to participate if required.

We cannot be completely certain of how this element of the study will run, as this will vary depending on the final tool. We have at this stage assumed that the tool will require contact at initial attendance and one early follow up visit (at around 5 days), and will include a mix of self-report and clinical tests. We have considered a range of other possibilities, and can provide these if needed.

Whilst we expect challenges, we do not anticipate any substantial difficulty with approach and recruitment.

Selection criteria:
Inclusion: Adults (>16 years of age) with an acute ankle sprain (<7 days old) who provide informed consent to participate. There will be no upper age limit.
Exclusions: Ankle fracture (excluding flake fracture < 2mm); Other recent (<3 months) lower limb fracture.

Procedure: ED staff will be asked to approach potential participants, provide a brief explanation of the study, administer the prognostic tool and record their clinical exam on a standard form. A standard approach will be instituted across all A&E departments based on our experience of running previous trials. We anticipate that completion of the tool will fall into the scope of any practitioner who is capable of taking a history and performing a simple ankle examination. In standard practice this would be nursing, medical, surgical or physiotherapy staff with experience or training in basic orthopaedic and injury management. We will develop a short pro-forma that serves two purposes (1) collection of routine core clinical data set in a tick box format and (2) A tick box to ensure that clinicians have provided potential participants with the trial information pack and a brief explanation of the trial. One copy of the form will be filed to the medical notes as a treatment record and a second copy is passed to the research team (with the patients consent) to notify them that a patient has been asked to participate in the trial. Provided that the clinical centres are involved in the development of the pro-forma, we have found this system to work well and once embedded in the system, to act as a reminder to invite patients into the trial. The form also enables us to check ED attendance statistics against the number of people being approached.

Formal consent to participate in the research will be taken by CLRN staff/the research team either by post or by telephone within one week. Participants will be consenting to allow us to use the data collected in the ED attendance (prognostic tool +/- any additional variables), and follow up questionnaires at 4 and 9 months so we can map the recovery trajectory and final recovery status at 9 months. A questionnaire at 4 months will also serve as reminder of the study, and as loss to follow up is likely to become larger over time, will ensure that we have responses on as many participants as possible.

Depending on the final format, the patient or ED clinicians (or both) will be asked to complete the prognostic tool. Although the predictor variables may be evident as per usual clinical practice, the prognosis will not be computed until final 9 month follow up is complete for each participant. Clinicians will be asked to continue to manage patients within the normal pathways at each of their hospitals, and we will record resource use. We may ask participants to return to the site, or complete some self-tests at home if further examination is indicated.

The questionnaires collected from participants at 4 and 9 months and will include details of severe and persistent symptoms, instability (as previous discussed), the validated Foot and Ankle Outcome Score,[39] health service resource use during follow up, and health related quality of life. Participants will be able to choose to respond via the web or a postal questionnaire. Additional demographic information will be collected if this is not included in the prognostic tool (for example sprain severity, age, gender, employment status). We are unlikely to include imaging techniques as either a component of the prognostic tool, or during follow up, although we will record if an image is taken and used in the practice. There is no indication in the literature that imaging adds to prognostic information, although a study of low dose CT scanning is being completed by one of the co-applicants of the study.

Sampling: The same principles apply to the sample size calculation stated here, as in Stage 1 justification for modelling. Current recommendations are that a minimum of 100 outcome events are required for external validation.[24,40] We anticipate the event rate to be between 26% and 32%, this would require an overall sample size of between 313 and 385. Assuming a 25% loss to follow-up and a lower event rate (20%) when recruiting all grades of ankles sprains we will target 675 patients to ensure the event rate is achieved.

Analysis: The performance of the tool will be assessed using calibration and discrimination metrics. Calibration will be assessed graphically with results for patients grouped by similar probabilities (tenths) and compare the mean predicted probability to the mean observed outcome.[23,24] The calibration plot will also be supplemented with estimates of the calibration slope and intercept. The discrimination of various prognostic models will be summarised with the concordance index (equivalent to the Area Under Receiver Operating Characteristic curve) with 95% confidence interval. Additional predictors not contained in the development dataset will be examined to see whether they improve the performance of the prognostic model using net reclassification improvement (NRI) and integrated discrimination improvement (IDI).[41]

Neither calibration nor discrimination captures the clinical usefulness of the prediction model (i.e. how does using the prediction model improve on a default policy of no prediction model). This will be examined using decision curve analysis, a relatively novel yet increasingly recommended approach in evaluating the predictive performance of a prediction model. It calculates the net benefit of using the model compared to not using the model,[32] evaluating and determining ranges of cut-off values for decision-making. The approach permits evaluation of the tools over a range of thresholds and the model with the greatest net benefit (.ie. the number of true positives penalised by a weighted proportion of the false-positives) will be considered as the preferred prediction model. The results will be converted into 'an additional x patients per 1000' who would be identified using the prediction tool compared to not using the tool. If additional predictors are found to be helpful, we will evaluate the incremental value of the predictors and update the model to accommodate these accordingly.[23,31] A model with the greatest net benefit (.ie. the number of true positives penalised by a weighted proportion of the false-positives) will be considered as the preferred prediction model. [32]

As part of the validation of the prognostic tool, we shall identify and evaluate the tool in pre-defined subgroups to evaluate the usefulness of the model in these subgroups. These subgroups will have different case-mix and thus if the model is shown to be useful in these subgroups, it increases the likelihood the model will be generalizable to other untested subgroups (e.g. primary care).[42]

Ethics: Ethical and governance approval will be sought from an MREC and National Institute for Health Research (NIHR) Coordinated System for gaining NHS Permission (CSP) via the Integrated Research Application System (IRAS). All participants will give informed consent to participate in the study, and we will seek permission for extended follow up if there is an appropriate data set to track patients through.

The initial approach will be made by a member of the ED clinical team. Verbal explanation of the study along with a study information leaflet will be given to all eligible patients. The study team will only know the identity of those patients that have agreed to participate. We will apply to the ethics committee to collect a minimum, anonymised data set on people declining (age, gender and severity of sprain) so that we can estimate generalizability. We will display posters in all participating departments to inform participants, along with the planned introduction from clinicians.

Estimated recruitment rate: We have scoped attendance at 7 EDs. The attendance rates are shown in Appendix 2. We are planning recruitment for 10 months and have assumed a conservative acceptance rate of 10-15%. Hence the average recruitment rate will be 15/ per centre/per month. We have scoped the CLRN arrangements at each site.

Stop go criteria: We use a staggered recruitment rate, but the viability of recruitment will be evident by 6 months of recruitment. If we are not recruiting at 70% of the proposed target or have not taken steps sufficient to remedy recruitment shortfall within the timetable, then the funder will be notified and closure considered.

Retention: Retention is challenging in acute injury trials. A reasonable proportion of the sample will be a mobile young population, and can be difficult to track. We will use methods developed in the HTA-funded CAST, BEST and MINT trials to ensure we maximise retention. We will use a small financial incentive in the form of vouchers at the 9 month follow up only.

2. Ethical considerations for all stages of the project

All aspects of the project will comply with Good Clinical Practice and the Declaration of Helsinki. All participants will be asked for written informed consent. Although some of the research is based in the Emergency Department with the attendant problems of that, the people we will approach will not fulfil the criteria of incapacity under the Mental Capacity Act (2005). NHSresearch ethics committees have previously approved a variety of telephone and postal consent methods recognising that this is an acute injury, that the ability to give a minimum of 24 hours between initial approach and consent may be compromised.

There are no ethical issues around collecting risk information without sharing this with clinicians or patients in the validation study. The rationale is that the prognostic tool is experimental until the end of the validation period. We do not know whether it enhances or detracts from routine clinical care. We will be very clear to ensure that all participating clinicians and patients are aware of this. We will not be withholding any form of standard practice. Neither do we shield the clinician from the responses made, only from the overall prediction score and interpretation. We will not calculate the overall prediction score until the end of the follow up period.

Risks and benefits:

The risks to participants are very small. This is an observational study, although we will be "intervening" with a prognostic tool. It is possible that the tool may include a performance test which if inappropriately implemented might cause prolongation or exacerbation of symptoms, or re-injury. We will act to minimise the risks of testing through careful selection of test items and training.

Data retention

Clinical Research Forms will be securely stored for 15 years after completion of the study. Electronic copies of the study data set will be retained indefinitely, and after a period of 3 years will be made available to other investigators on request and subject to a data sharing agreement. Oxford Clinical Trials Research Unit (OCTRU) has Standard Operating Procedures to ensure adherence with the Data Protection Act (2003), ICH and MRC good clinical practice guidelines.

Monitoring and safety reporting

The trial will have a Trial Steering Committee (comprising 3 independent members and 2 dependent members (Lamb and Williams). We will not formally constitute a separate Data Monitoring and

Ethics Committee as this is not a comparative interventional trial. The three independent committee members will review safety, recruitment and ethical data, and would be able to close the trial.

We will use a risk adapted approach to monitoring. Recruitment and progress against milestones will be monitored by the study team at monthly Trial Management Meetings. A quality assurance programme will ensure consistent and high quality approaches to consent and implementation of the clinical rule. We will define an adverse event as one that occurs directly as a result of the clinical prediction tool (patient report of substantial worsening of pain and/or symptoms within 2 days of the test) and serious adverse events as any hospitalisation or death that occurs within 2 days of the test, and would not have been anticipated as part of clinical management at initial presentation.

Inclusion of people from ethnic minorities

Our sampling frame of EDs include a number of departments that have catchments with substantial ethnic minority communities (in particular Birmingham Heartlands). In the HTA funded Managing Injuries of the Neck Trial, we translated materials into a range of languages but found no evidence that the translated materials were used in preference to an English language version. We will explore the need to translate materials at each site, as well as the range of languages and take appropriate action. Translation services are available in NHS Trusts.

3. Research Governance

The research will be sponsored by the University of Oxford (with appropriate insurance policies in place) and follow Oxford Clinical Trial Research Units Standard Operating Procedures. This study will be fully compliant with the research governance framework and MRC Good Clinical Practice guidelines.

4. Study Management

This study will require careful management with stages 1 and 2 feeding vital information into stage 3. The Principal Investigator will be Professor Lamb. Day-to-day management will be undertaken by Dr Williams and an experienced Study Co-ordinator. Study management group meetings will be held on a monthly basis to monitor progress and obtain input from the appropriate co-applicants and site investigators at the relevant stages of the study. In addition, 6-monthly co-applicant meetings will be held to ensure close involvement from all parties.

5. Project Timetables and Milestones

We propose an ambitious two and a half year study beginning in November 2014 (Table 2). We will begin governance and approval processes in the pre-funding stage. We will analyse the CAST data set and systematically review the literature concurrently over the first 3 months. Immediately following this we will conduct the expert and PPI consensus process in order that the initial tool is finalised ready for validation at the end of month 6. We will have approvals in place to commence recruitment in month 7 with the 5 site initiations staggered over 2 months. Follow-up of participants will continue until month 25. Data cleaning, analysis and final reporting will be carried out over the last 6 months of the project.

Table 2 – Project timetable

Quarter	1	2	3	4	5	6	7	8	9	10
Stage 1										
Analysis of CAST data set										
Stage 2										
Systematic Review										
Expert and PPI consultation										
Prognostic tool finalised for validation										
Stage 3										
Ethical and governance approvals										
Site training and set up										
Recruitment for ED cohort study										
Follow-up for ED cohort study										
Data entry and cleaning										
Analysis, reporting and dissemination										

6. Research Team

Professor Sallie Lamb will be Principal Investigator. She is a physiotherapist and a fellow of the Royal Statistical Society. She has been CI on NIHR HTA-funded studies that span primary, secondary and tertiary care, including trials in NHS EDs.
Dr Gary Collins will lead the statistical input. He is currently the Principal Investigator of an MRC methodology grant to develop statistical methods and reporting guidelines for prognostic studies.
Dr Mark Williams will be responsible for the day to day management of the study. He previously co-ordinated the HTA funded Managing Injuries of the Neck Trial which recruited over 3000 patients through EDs in 18 months. He contributed to recruitment in the HTA funded CAST trial.
Prof Steve Goodacre will provide ED clinical and research expertise and act as local PI for recruitment in the validation cohort study. He has been PI and co-applicant on numerous NIHR HTA-funded studies in EDs including RATPAC (n=2243) and 3Mg (n=1109) trials.
Prof Matthew Cooke will also provide ED clinical and research expertise and act as local PI for recruitment in the validation cohort study at Heart of England Foundation Trust. He was previously National Clinical Director for Urgent and Emergency Care.
Mr Steve Gwilym will provide expertise input as an academic trauma surgeon with a special interest in pain.
Dr Phil Hormbrey will provide ED clinical and research expertise and act as the local PI in Oxford.
Dr David Wilson will provide radiology clinical and research expertise. He is currently conducting a pilot prospective cohort study to evaluate prognostic factors for poor outcome for ankle sprains, with a particular focus on the role of imaging (REC Ref 12/SC/0596).
Ms Jennifer Bostock is the lead PPI representative. She has experience as an NIHR lay reviewer, PPI representative and has served for many years on a Research Ethics Committee.

Collaborators: We have agreement from 10 NHS trusts with named local PIs willing to participate and also to contribute to the refinement of the prognostic tool in Stage 2.

7. Patient and Public Involvement

PPI input has been co-ordinated with the support of the South Central Research Design Service. To date, we have recruited 4 PPI representatives from a process of open advert on the People in Research website, South Central RDS e-bulletin, and the John Radcliffe Hospital Emergency Department in Oxford. Our appointed PPI representatives have experienced an ankle sprain and accessed NHS ED services. Ms Jennifer Bostock has agreed to be the PPI lead representative and has experience as an NIHR lay reviewer and has served for many years on a Research Ethics Committee.

In order to develop and refine our application, we have held a programme development meeting with our PPI representatives. This initial meeting led to further discussion and agreement to collaborate. Our representatives have reviewed and contributed to ideas and have provided feedback on our proposed programmes of work including who the team should consist of; the experience of service use from the PPI perspective; the relevance of our proposed outcomes; acceptability of the research methods and the role of PPI input in developing and guiding the full application and research programme. We have sought input on what are important outcomes and these have influenced the make-up of our composite outcome measure. They have commented on this application form, reviewers' comments and the according adaptations.

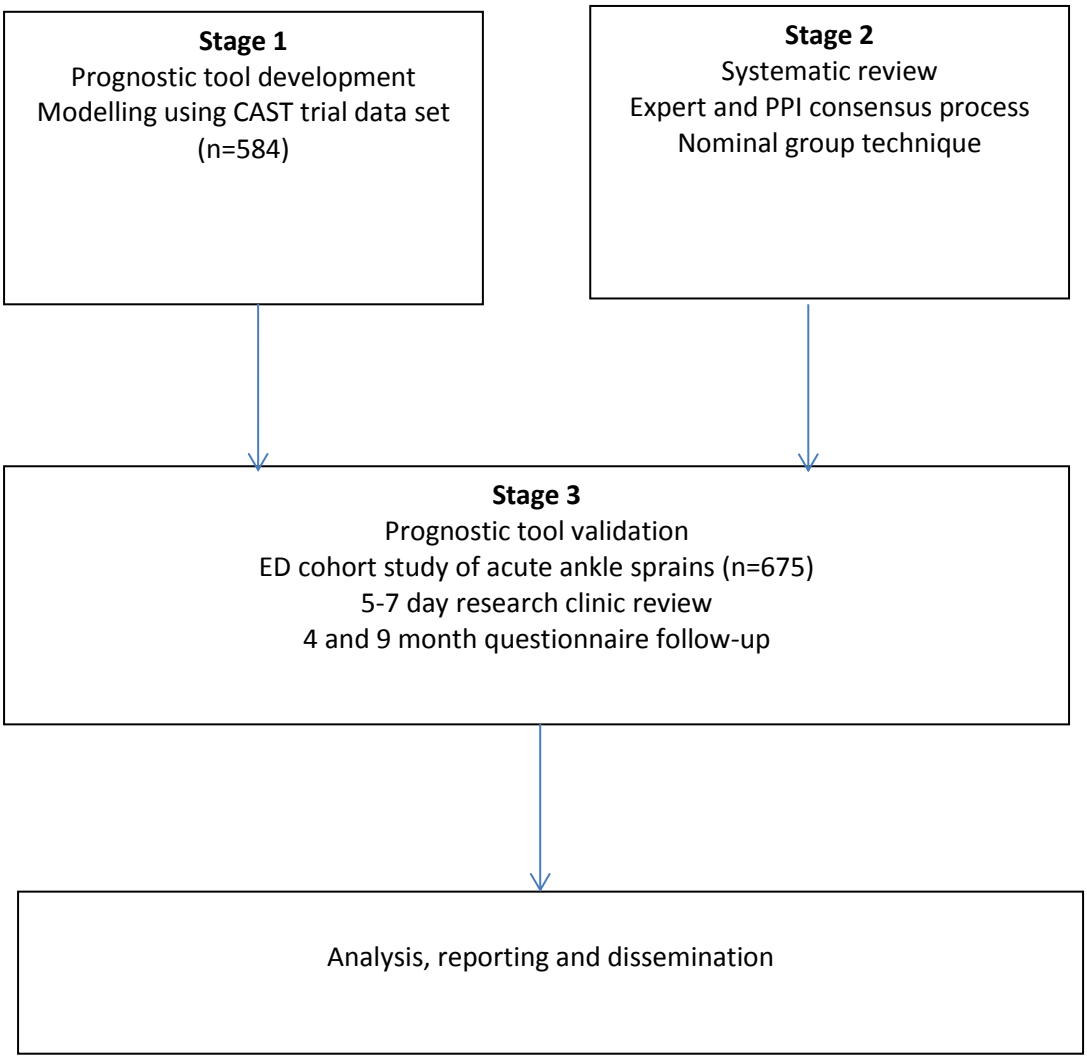
Stage 2: We will consult our PPI reps further regarding design of the study and important issues regarding outcome measures.

Stage 3: Our experience has shown the value of early PPI input into the design of patient-facing materials to ensure they are understandable and address issues of relevance to ED patients. We will seek input on study procedures to ensure that burdens placed on participants are justifiable and well explained, also to identify potential recruitment barriers.
Reporting & dissemination: We will collaborate with our reps when writing the funders' report, subsequent publications and developing a strategy for dissemination to users. Attendance at conferences by PPI reps will be supported (e.g. INVOLVE conference). Training needs for reps will be addressed on an individual basis. Management: PPI reps will be invited to attend 6 monthly collaborator meetings. A PPI rep will appointed to the Independent Steering Committee.

8. Dissemination and projected outputs

We will publish an HTA monograph of the project. Currently no systematic review has been published specifically investigating prognostic factors for poor recovery following acute ankle sprain. We will publish a systematic review of prognostic factors in a peer-reviewed journal. We anticipate also publishing the results of the modified nominal group element of the study, which is likely to also yield new perspectives on the future research agenda for ankle sprain. Following analysis of the internal and external validation of the tool, we will publish the findings in a peer-reviewed journal and disseminate through international conferences. Any publications will adhere to the PRISMA (for systematic review)[43] and STROBE guidelines[44]. We also plan to disseminate to patient and public groups e.g INVOLVE conference. More recently our research group has been looking to social media to promulgate research findings. We are establishing a monthly Blog and Twitter and will utilise these media to increase visibility of the research findings. Our group is working alongside the Oxford CLARHC, so we would hope to utilise this infra-structure to study and promote implementation. Should we develop a suitable tool, we will make the tool freely available on the web, with a simple registration to track areas and level of impact. We will investigate the possibility of using a mobile app for scoring the tool and utilising existing infrastructure used by clinicians to access decision tools e.g. MDCalc.com

9. Flow diagram



10. Appendices

Appendix 1 - List of variables available from CAST dataset related to clinical and empirical importance

	Available in CAST data set	Not available in CAST data set
Clinically important	Grade of sprain Weight Bearing ability Pain intensity Swelling (S-R) Initial Ligament testing Recurrent sprains Reduced Range of Motion (S-R) Type of employment Type of recreational activities Frequency of sports participation	Reduced Range of Motion (C-R) Location of injury Syndesmosis test positive
Moderate to strong relationship with outcome (function unless specified)	Previous sprain (outcome pain/symptoms)[1] Males[1] C-R severity[2] Palpation score[2] Function score[2] Pain[2] Swelling[2] Ankle Function Score[2,3] ED attendance (outcome instability)[3] S-R activity limitation[4] Age[5,6] Weight-bearing status[5] Injury mechanism[5] Females[6]	High level athlete[1] Syndesmosis sprain[7]
Items confirmed in prognostic tool	None	None
Abbreviations	S-R = Self-Reported C-R = Clinician-Reported	
References	1. Linde (1986), 2. de Bie (1997), 3. van Middelkoop (2012), 4. Wilson and Gansneder (2000), 5. O'Connor et al (2013), 6. Akacha and Hutton (2011), 7. Gerber et al (1998)	

Appendix 2 - Sites for Stage 3 validation cohort

Sites/Trust	Local Investigator	Number of adult sprains per annum/month	Estimated number of eligible patients recruited per month	Current CLRN support presence in ED
Heart of England NHS Foundation Trust	Prof Matthew Cooke	2387 per annum	15-20	N but CLRN-funded CRF
Sheffield Teaching Hospitals NHS Foundation Trust	Prof Steve Goodacre	2123 per annum	15-20	Y
Oxford University Hospitals NHS Trust	Dr Phil Hormbrey	70-130 per month	7-13	Y
Gloucestershire Hospitals NHS Foundation Trust	Dr Vicky Stacey	CGH - 787 per annum; GRH 1019 per annum	13-16	N
Worcestershire Acute NHS trust	Mr Richard Morrell	720 per annum	8-10	N
South Warwickshire NHS Foundation Trust	Mr Matthew Dunn	1359 per annum	10-15	N
University Hospitals Leicester NHS Trust	Prof Tim Coats	120-140 per month	12-14	Y
South Tees NHS Trust	Dr Paul Hunt	100-120 per month	10-12	N

- Additional sites that have agreed to participate but unable to provide data are:

Walsgrave Hospital / University Hospitals Coventry and Warwickshire NHS Trust, Frenchay Hospital / North Bristol NHS Trust

11. References

1 Cooke, M. *et al.* Treatment of severe ankle sprain: a pragmatic randomised controlled trial comparing the clinical effectiveness and cost-effectiveness of three types of mechanical ankle support with tubular bandage. The CAST trial. *Health Technol Assess* **13**, iii, ix-x, 1-121, doi:10.3310/hta13130 (2009).

2 van Rijn, R. M. *et al.* What is the clinical course of acute ankle sprains? A systematic literature review. *Am J Med* **121**, 324-331.e326, doi:10.1016/j.amjmed.2007.11.018 (2008).

3 Polzer, H. *et al.* Diagnosis and treatment of acute ankle injuries: development of an evidence-based algorithm. *Orthop Rev (Pavia)* **4**, e5, doi:10.4081/or.2012.e5 (2012).

4 Wikstrom, E. A., Hubbard-Turner, T. & McKeon, P. O. Understanding and treating lateral ankle sprains and their consequences: a constraints-based approach. *Sports Med* **43**, 385-393, doi:10.1007/s40279-013-0043-z (2013).

5 Verhagen, R. A., de Keizer, G. & van Dijk, C. N. Long-term follow-up of inversion trauma of the ankle. *Arch Orthop Trauma Surg* **114**, 92-96 (1995).

6 Stiell, I. *et al.* Multicentre trial to introduce the Ottawa ankle rules for use of radiography in acute ankle injuries. Multicentre Ankle Rule Study Group. *BMJ* **311**, 594-597 (1995).

7 van Dijk, C. N., Mol, B. W., Lim, L. S., Marti, R. K. & Bossuyt, P. M. Diagnosis of ligament rupture of the ankle joint. Physical examination, arthrography, stress radiography and sonography compared in 160 patients after inversion trauma. *Acta Orthop Scand* **67**, 566-570 (1996).

8 Cooke, M. W., Lamb, S. E., Marsh, J. & Dale, J. A survey of current consultant practice of treatment of severe ankle sprains in emergency departments in the United Kingdom. *Emerg Med J* **20**, 505-507 (2003).

9 Hiller, C. E. *et al.* Characteristics of people with recurrent ankle sprains: a systematic review with meta-analysis. *Br J Sports Med* **45**, 660-672, doi:10.1136/bjsm.2010.077404 (2011).

10 Linde, F., Hvass, I., Jürgensen, U. & Madsen, F. Early mobilizing treatment in lateral ankle sprains. Course and risk factors for chronic painful or function-limiting ankle. *Scand J Rehabil Med* **18**, 17-21 (1986).

11 de Bie, R. A., de Vet, H. C., van den Wildenberg, F. A., Lenssen, T. & Knipschild, P. G. The prognosis of ankle sprains. *Int J Sports Med* **18**, 285-289, doi:10.1055/s-2007-972635 (1997).

12 Wilson, R. W. & Gansneder, B. M. Measures of functional limitation as predictors of disablement in athletes with acute ankle sprains. *J Orthop Sports Phys Ther* **30**, 528-535, doi:10.2519/jospt.2000.30.9.528 (2000).

13 van Middelkoop, M., van Rijn, R. M., Verhaar, J. A., Koes, B. W. & Bierma-Zeinstra, S. M. Re-sprains during the first 3 months after initial ankle sprain are related to incomplete recovery: an observational study. *J Physiother* **58**, 181-188, doi:10.1016/S1836-9553(12)70109-1 (2012).

14 Gerber, J. P., Williams, G. N., Scoville, C. R., Arciero, R. A. & Taylor, D. C. Persistent disability associated with ankle sprains: a prospective examination of an athletic population. *Foot Ankle Int* **19**, 653-660 (1998).

15 O'Connor, S. R., Bleakley, C. M., Tully, M. A. & McDonough, S. M. Predicting functional recovery after acute ankle sprain. *PLoS One* **8**, e72124, doi:10.1371/journal.pone.0072124 (2013).
16 Akacha, M. & Hutton, J. L. Modelling the rate of change in a longitudinal study with missing data, adjusting for contact attempts. *Stat Med* **30**, 1072-1089, doi:10.1002/sim.4165 (2011).
17 Kerkhoffs, G. M. *et al.* Diagnosis, treatment and prevention of ankle sprains: an evidence-based clinical guideline. *Br J Sports Med* **46**, 854-860, doi:10.1136/bjsports-2011-090490 (2012).
18 Kerkhoffs, G. M. *et al.* Immobilisation and functional treatment for acute lateral ankle ligament injuries in adults. *Cochrane Database Syst Rev*, CD003762, doi:10.1002/14651858.CD003762 (2002).
19 Kerkhoffs, G. M., Handoll, H. H., de Bie, R., Rowe, B. H. & Struijs, P. A. Surgical versus conservative treatment for acute injuries of the lateral ligament complex of the ankle in adults. *Cochrane Database Syst Rev*, CD000380, doi:10.1002/14651858.CD000380.pub2 (2007).
20 Lamb, S. E. *et al.* Mechanical supports for acute, severe ankle sprain: a pragmatic, multicentre, randomised controlled trial. *Lancet* **373**, 575-581, doi:10.1016/S0140-6736(09)60206-3 (2009).
21 Moons, K. G., Royston, P., Vergouwe, Y., Grobbee, D. E. & Altman, D. G. Prognosis and prognostic research: what, why, and how? *BMJ* **338**, b375 (2009).
22 Bleakley, C. M. *et al.* Effect of accelerated rehabilitation on function after ankle sprain: randomised controlled trial. *BMJ* **340**, c1964 (2010).
23 Harrell, F. E. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. (Springer, 2001).
24 Steyerberg, E. W. *Clinical prediction models: a practical approach to development, validation, and updating*. (Springer, 2009).
25 Peduzzi, P., Concato, J., Kemper, E., Holford, T. R. & Feinstein, A. R. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* **49**, 1373-1379 (1996).
26 Vittinghoff, E. & McCulloch, C. E. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol* **165**, 710-718, doi:10.1093/aje/kwk052 (2007).
27 Royston, P., Altman, D. G. & Sauerbrei, W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* **25**, 127-141, doi:10.1002/sim.2331 (2006).
28 Royston, P. & Sauerbrei, W. Building multivariable regression models with continuous covariates in clinical epidemiology--with an emphasis on fractional polynomials. *Methods Inf Med* **44**, 561-571 (2005).
29 Vergouwe, Y., Royston, P., Moons, K. G. & Altman, D. G. Development and validation of a prediction model with missing predictor data: a practical approach. *J Clin Epidemiol* **63**, 205-214, doi:10.1016/j.jclinepi.2009.03.017 (2010).
30 van Rijn, R. M., Willemsen, S. P., Verhagen, A. P., Koes, B. W. & Bierma-Zeinstra, S. M. Explanatory variables for adult patients' self-reported recovery after acute lateral ankle sprain. *Phys Ther* **91**, 77-84, doi:10.2522/ptj.20090420 (2011).
31 Sullivan, L. M., Massaro, J. M. & D'Agostino, R. B. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Stat Med* **23**, 1631-1660, doi:10.1002/sim.1742 (2004).

32 Vickers, A. J. & Elkin, E. B. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* **26**, 565-574, doi:10.1177/0272989X06295361 (2006).

33 Altman, D. G. in *Systematic Reviews in Health Care, Meta-analysis in context* (eds M. Egger, G. Davey Smith, & D. G. Altman) Ch. 13, (BMJ publishing group, 2001).

34 Huguet, A. *et al.* Judging the quality of evidence in reviews of prognostic factor research: adapting the GRADE framework. *Syst Rev* **2**, 71, doi:10.1186/2046-4053-2-71 (2013).

35 Hayden, J. A., van der Windt, D. A., Cartwright, J. L., Côté, P. & Bombardier, C. Assessing bias in studies of prognostic factors. *Ann Intern Med* **158**, 280-286, doi:10.7326/0003-4819-158-4-201302190-00009 (2013).

36 Hayden, J. A., Côté, P., Steenstra, I. A., Bombardier, C. & Group, Q.-L. W. Identifying phases of investigation helps planning, appraising, and applying the results of explanatory prognosis studies. *J Clin Epidemiol* **61**, 552-560, doi:10.1016/j.jclinepi.2007.08.005 (2008).

37 WILLIAMSON, E., WILLIAMS, M., GATES, S. & LAMB, S. A systematic literature review of psychological factors and the development of late whiplash syndrome. *Pain* **135**, 20-30, doi:10.1016/j.pain.2007.04.035 (2008).

38 Black, N., Brazier, J. & Fitzpatrick, R. *Health services research methods: a guide to best practice*. (BMJ Books, 1998).

39 Roos, E. M., Brandsson, S. & Karlsson, J. Validation of the foot and ankle outcome score for ankle ligament reconstruction. *Foot & Ankle International* **22**, 788-794 (2001).

40 Vergouwe, Y., Steyerberg, E. W., Eijkemans, M. J. & Habbema, J. D. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol* **58**, 475-483, doi:10.1016/j.jclinepi.2004.06.017 (2005).

41 Pencina, M. J., D'Agostino, R. B. & Vasan, R. S. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* **27**, 157-172; discussion 207-112, doi:10.1002/sim.2929 (2008).

42 Altman, D. G. & Royston, P. What do we mean by validating a prognostic model? *Stat Med* **19**, 453-473 (2000).

43 Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* **339**, b2535 (2009).

44 von Elm, E. *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* **61**, 344-349, doi:S0895-4356(07)00436-2 [pii] 10.1016/j.jclinepi.2007.11.008 (2008).