Title
Prognostic models in adults undergoing physiotherapy for rotator cuff disorders - a systematic review

Systematic review registration number (PROSPERO)
CRD42014008973

Authors
1. Cordula Braun\textsuperscript{1,2}, MSc
2. Nigel C. Hanchard\textsuperscript{1}, PhD
3. Alan M. Batterham\textsuperscript{1}, PhD
4. Helen H. Handoll\textsuperscript{1}, DPhil
5. Andreas Betthäuser\textsuperscript{3}, Dr med

Affiliations
\textsuperscript{1}Health and Social Care Institute, School of Health and Social Care, Teesside University, Middlesbrough, United Kingdom;
\textsuperscript{2}Department of Health (Physiotherapy), hochschule 21, Buxtehude, Germany;
\textsuperscript{3}Schulter-Zentrum Hamburg, Hamburg, Germany

Corresponding author
Cordula Braun, c/o hochschule 21, Department of Health (Physiotherapy), Harburger Str. 6, 21614 Buxtehude, Germany; Telephone: 0049-1713024405 (mobile); email: C.Braun@tees.ac.uk.
Prognostic models in adults undergoing physiotherapy for rotator cuff disorders - a systematic review

Abstract

Background

Rotator cuff related disorders represent the largest subgroup of shoulder complaints. Despite the availability of various conservative and surgical treatment options, the precise indications for these options remain unclear.

Purpose

The purpose of this systematic review was to synthesize the available research on prognostic models for predicting outcomes in adults undergoing physiotherapy for painful rotator cuff disorders.

Data sources

We searched Medline, Embase, Cinahl, Cochrane CENTRAL, PEDro and trial registries up to October 2015.

Study selection

We included primary studies exploring prognostic models in adults undergoing physiotherapy, with or without other conservative measures, for painful rotator cuff disorders. Primary outcomes were pain, disability and adverse events. Inclusion was limited to prospective investigations of prognostic factors elicited at the baseline assessment. Study selection was independently performed by two reviewers.
We used a piloted form to extract data on key aspects of study design, characteristics, analyses and results. Risk of bias and applicability was independently assessed by two reviewers using the PROBAST tool.

Five studies were included in the review. These were extremely heterogeneous in many aspects of design, conduct and analysis. The findings were analysed narratively.

All included studies were rated as at high risk of bias, and none of the resulting prognostic models was found to be usable in clinical practice.

There are no prognostic models ready to inform clinical practice in the context of the review question, highlighting the need for further research on prognostic models for predicting outcomes in adults who undergo physiotherapy for painful rotator cuff disorders. The design and conduct of future studies should be receptive to developing methodologies.

Word count of body of article (unmasked copy)

4,384
Introduction

Painful shoulder complaints are among the commonest musculoskeletal disorders in medical and physiotherapy practice. These may become persistent, potentially leading to increased use of healthcare resources and prolonged sick leave, and placing a cost burden on the individual and society. Most shoulder complaints (29% to 85%) involve the subacromial-subdeltoid bursa and rotator cuff. The pathology is diverse, reflecting a degenerative continuum from tendinopathy to partial (PTT) or full-thickness tears (FTT). Rotator cuff tears, in particular, have a reported prevalence of over 40% in symptomatic shoulder pain populations and are strongly correlated with age. Clinical features of rotator cuff disorders may include pain, abnormalities on tests of rotator cuff function and integrity, and significantly impaired shoulder function and health-related quality of life. While diagnosis of rotator cuff disorders is based on clinical signs and symptoms, verification of a rotator cuff tear requires diagnostic imaging (e.g. ultrasonography, magnetic resonance imaging).

Initial treatment of rotator cuff disorders usually involves medical care (e.g. oral medication, corticosteroid injections) and physiotherapy (e.g. exercises, manual therapy). Current guidelines advise conservative treatment as the first-line treatment, with surgery mainly reserved for non-responders. Direct comparisons of conservative versus surgical treatment have not shown clinically relevant differences between groups. Nonetheless, the rates of surgical intervention for rotator cuff disease have considerably increased in many countries. Unnecessary surgery is undesirable, as is ineffective conservative treatment. Patients and health care providers alike would benefit if likely responders and, by corollary, non-responders to conservative interventions, could be identified at the commencement of the care pathway. This would avoid unnecessary suffering, reduce uncertainty and
anxiety and limit exposure to the risks of surgery, as well as conserving limited resources. “Understanding which patients [with rotator cuff tears] do best with non-operative treatment” has been rated a top “priority scientific research issue”.

The importance of predicting which patients will respond to particular treatments is increasingly recognised and has stimulated interest in prognosis and prognosis research. There has been a corresponding development in prognosis research methodology. Prognosis research aims to predict clinical outcomes in individual patients. One aspect of prognosis research involves single factors, which, in the context of painful rotator cuff disorders, would typically be demographic or clinical. However, single factors are unlikely to predict outcomes satisfactorily. Multivariable prognostic models are better placed to do so, because they account for real-life clinical complexities. An illustration of a multivariable model is the Nottingham Prognostic Index (NPI), which is used to predict survival of women diagnosed with primary breast cancer by the following formula: $\text{NPI} = (0.2 \times \text{tumor diameter (cm)}) + \text{lymph node stage} + \text{tumour grade}$. Scores are interpreted by reference to a table.

Prognostic modelling encompasses three key phases: development (including internal validation, i.e. determining the model’s replicability using data from the primary sample); external validation (determining the model’s generalizability using data from independent samples); and investigation of clinical impact (a model’s effectiveness and cost-effectiveness in improving outcomes). External validation is a crucial step before a model can be considered usable in clinical practice. The objective of this review was to synthesize the available research on prognostic models for predicting outcomes in adults who undergo physiotherapy for painful...
rotator cuff disorders. We aimed to provide a resource to facilitate clinical decision-making but also to identify any research gaps. To our knowledge, this is the first systematic review to synthesize the available evidence on this topic.

Methods

Overall approach

We based our methods on the recent recommendations of the PROGRESS (PROgnosis RESearch Strategy) partnership and, complementarily, the Cochrane Prognosis Methods Group. We used PROGRESS terminology where possible. This review is based on an a priori protocol, registered in PROSPERO, the International Prospective Register of Systematic Reviews (registration nr. CRD42014008973), and available at www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014008973#.VTodb mYom1k. Differences between protocol and review are specified within the supplementary material (Table A.1).

Criteria for considering studies for inclusion

Types of studies

We included primary studies exploring prognostic models for predicting outcomes in adults undergoing physiotherapy, with or without other conservative measures, for painful rotator cuff disorders. Inclusion encompassed any of the three phases of prognostic research. We considered any prospective longitudinal research designs. There was no language restriction on searches. Only reports written in English were included, but we planned to document relevant studies reported in other languages.
This review addressed adults (age ≥ 18 years) diagnosed with painful rotator cuff disorders, at any stage, which was unrelated to substantial trauma (e.g. dislocation). We placed no restriction on how this was diagnosed. We also included studies whose inclusion criteria were symptoms or mechanisms consistent with rotator cuff disorders, e.g. “subacromial pain”, “subacromial impingement” or “shoulder impingement”. Studies in which 85% or more of participants satisfied our criteria were included. We did not actively seek studies focussed on subacromial–subdeltoid bursitis, although, due to its intimate relationship with the rotator cuff, incidental involvement of this bursa may well occur in our population of interest. There was no restriction on the duration or severity of symptoms at baseline, or on the care setting.

We excluded studies focusing on people who were pain-free or had trauma-related conditions, and studies on calcific tendinitis or disorders of the long head of biceps. We anticipated that in some studies there would be insufficient characterisation of participants (e.g. that other potential causes of shoulder pain might not be considered). In these cases, we erred on the side of inclusivity.

Interventions

We included studies evaluating physiotherapy, of any duration or frequency, with or without other conservative measures as part of a non-surgical care pathway. Physiotherapy had to involve therapeutic exercises and/or manual techniques, as these are considered the core interventions, but could include adjunctive treatments (e.g. acupuncture, electrotherapy, corticosteroid injections, osteopathic musculoskeletal interventions or thermotherapy). Studies comparing physiotherapy
versus a non-physiotherapy control group were only considered if there was separate
prognostic modelling for the former.

Prognostic factors
For simplicity, we applied the term “prognostic factor” to any factor under
investigation, regardless of whether it was (or had previously been) found to have
prognostic properties. We required these factors to be elicited at the baseline
assessment.

Outcomes
Primary outcomes were

- Pain
- Shoulder disability on a validated patient-reported outcome measure (PROM),
e.g. Oxford Shoulder Score
- Adverse events (e.g. exacerbations of symptoms)

Secondary outcomes were

- Health-related quality of life (HrQoL), e.g. Short Form 36 (SF-36)
- Sick leave
- Patient’s global perception of change (GPC)
- Imaging determination of structural progression of tear
- Patient’s decision to undergo surgery

To be included, a study had to present a prognostic model in relation to at least one
of these outcomes.

Types of analysis
Studies had to evaluate prognostic models of multiple factors, but no restriction was placed on the phase of research or on the type of multivariable analysis. Furthermore, the models had to be presented in full in the study report or provided on request by the corresponding authors.

Data sources and searches

Electronic searches

Building on the experience of previous searches for a prognostic study (2011-12, report in preparation) and two systematic reviews of interventions in this field,\textsuperscript{32,33} we developed a broad strategy including only search terms relating to the population and interventions. For Medline, we used a slightly amended version of a filter developed for prognosis research;\textsuperscript{34} see Table A.2 for the full search strategy.

We searched the following electronic databases from inception: Medline (EBSCO), Embase (Ovid), Cochrane CENTRAL (Ovid), Cinahl (EBSCO), PEDro and The World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP).

The formal database search was initially run on 16 May 2014 (ICTRP was searched on 14 Aug) and updated to 19 October 2015. One author (CB) conducted the searches. We followed up “related articles” suggestions for all relevant studies.

Searching other sources

We supplemented the electronic searches by hand searching the reference lists of all relevant studies and existing prognosis systematic reviews on shoulder pain. We further matched the compilation of eligible studies with the results from our previous searches.
Study selection

Study selection was independently performed by two authors (CB and NCH or CB and HHH). In case of disagreement, consensus was sought through discussion or involvement of a third person (AMB, HHH).

Data extraction and quality assessment

Data extraction and management

We used a purpose-designed and piloted form to extract data on key aspects of study design, characteristics, analyses and results. For developmental studies, we extracted only one model per study: either the reportedly final model or the most complete model including the main effects for all prognostic factors. We extracted key statistics of the models and of model performance as reported by the studies. Extraction of summary statistics of predictive performance, where possible, included the standard error of the estimate (SEE) for studies with continuous outcomes and likelihood ratios or area under the curve (c statistics) for studies with binary outcomes. We also reported any further measures of model performance (e.g. of the model’s discriminative ability), and validation (internal or external). Two authors (CB, NCH) independently extracted the data. We did not impute missing data. We limited author contact to the clarification of issues related to study eligibility.

Assessment of risk of bias and applicability

To assess risk of bias and applicability, we used the latest available version of the PROBAST tool (Prediction Study Risk of Bias Assessment Tool), which at the time of writing was in the late stages of development but unpublished (personal communication). PROBAST is designed to assess risk of bias and applicability of primary studies evaluating (developing and/or validating) prognostic models. It is
domain-based, with a similar structure to QUADAS-2. It has five key domains: participant selection, predictors (i.e. prognostic factors), outcome, sample size and participant flow, and analysis. Each domain comprises a set of “signalling questions” to facilitate judgements about risk of bias: low, high or unclear. Additionally, the first three domains are assessed for concerns (low, high or unclear) about the applicability of the study’s design and characteristics to the review question. A summative judgement across all domains leads to an overall rating of low, high or unclear risk of bias or concern about applicability. Lastly, the usability of the model is rated as yes or no. For this item, we considered whether the model was ready for use in the intended context and target population, in view of the phase of research, the detail with which the model was presented, and the risk of bias. Risk of bias and applicability assessment was independently performed at study level by two authors (CB, NCH). In case of disagreement, consensus was sought through discussion or involvement of a third person (AMB or HHH).

Data synthesis and analysis

All included studies were tabulated and narratively synthesised. In the absence of sufficient good quality, comparable and externally validated studies, we did not undertake quantitative data synthesis.

Results

Search and selection process

Figure 1 outlines the complete process. The titles and abstracts of 5,889 results overall were screened. Fifty-four full text articles were obtained and considered for inclusion, six of which were identified from previous prognosis systematic reviews, five from our previous searches, and one from personal
communication.\textsuperscript{39} We included five studies\textsuperscript{39-43} and excluded 49 (see Table A.3 for further details). The most frequent reason for exclusion was a lack of multivariable prognostic modelling. We identified (by protocol or registry entries) eight clearly or potentially relevant ongoing studies (see Figure 1 and Table A.4). We obtained unpublished full multivariable model data relating to the trial of Björnsson Hallgren et al.\textsuperscript{40}

\textit{Included studies}

Study characteristics

Key characteristics of the five studies are presented in Table 1. The studies were published between 2005 and 2014. All appeared to have been conducted in outpatient settings. All were cohort studies, but in two the cohort was derived from pooled data from an RCT.\textsuperscript{39,40} None of the studies was prospectively registered; however, the intention for a prognostic investigation was mentioned in the published protocol\textsuperscript{44} for the study by Kromer et al. 2014. Four studies concerned model development and the fifth\textsuperscript{42} was reported as a validation study.

Four studies\textsuperscript{39-41,43} investigated mixed populations with impingement-related shoulder pain. One of these\textsuperscript{41} excluded FTT. One study\textsuperscript{42} investigated a rotator cuff tear population without differentiating between PTT and FTT. Initial sample sizes ranged from 33\textsuperscript{41} to 102;\textsuperscript{40,43} with the number of outcome events (number of patients in whom the prognosticated event occurred) ranging from 23\textsuperscript{41} to 89.\textsuperscript{43} Although varying in duration, content and dosage, physiotherapy was provided to all study participants; steroid injections were provided to all participants of one study\textsuperscript{40} and were optional in another.\textsuperscript{43}
The number of initially considered prognostic factors was unclear in three studies\textsuperscript{39-41} but, based on the presented data, appeared to range from eight\textsuperscript{43} to presumably over 60.\textsuperscript{41} Prognostic factors mainly involved demographics and clinical characteristics such as symptoms or diagnostic imaging findings. One study\textsuperscript{39} investigated psychosocial factors. None of the studies provided a full and unambiguous rationale for all initially considered factors. Though, in some cases, reference was to previous prognosis research, the approaches to the literature appeared non-systematic. Kromer et al.\textsuperscript{39} presented some focussed, literature-based justification for two of the factors modelled: fear avoidance beliefs and catastrophizing. Apart from these two exceptions, prognostic factors were not systematically derived from the literature.\textsuperscript{39-43}

Each study used different outcome measures, but all included PROMS; the outcomes used for this review are presented in Table 1. Follow-up ranged from six weeks\textsuperscript{41} to 12 months.\textsuperscript{39,40,42,43}

The methods for selecting prognostic factors for multivariable analysis, where specified, varied (Table A.5); two studies\textsuperscript{39,41} explicitly reported using some automated statistical method, e.g. analysis of univariable correlations between the prognostic factors and the outcome.

Approaches to multivariable modelling also varied. An automated statistical process, e.g. stepwise regression, was used in three studies.\textsuperscript{39,41,43} The nominal validation study by Merolla\textsuperscript{42} was severely flawed by inappropriate statistical analysis.

Risk of bias and applicability
Table 2 presents the summary of our PROBAST assessment. All studies were overall rated to be at high risk of bias; this was mainly due to issues within domains 3 to 5 (outcome, sample size and flow, and analysis). Ratings were affected by numerous issues, namely: inclusion of prognostic factors in the outcome definition; unclear or lack of blinding of outcome determination to prognostic factor information; an unreasonable number (> 5) of prognostic factors in relation to the number of outcome events (which we assessed in relation to the number of factors included reportedly final model or, where this was not specified, the most complete model including main effects for all prognostic factors); unclear handling of missing data; use of univariable analyses to select prognostic factors; unclear or unspecified modelling methods; and failure to consider overfitting of data, complexities in the data, evaluation of performance measures or non-linear relationships.

Overall concerns about applicability mainly related to domain 2 (predictors) and were rated as low for two studies, unclear for one, and high for two. The high concerns related to the definition and assessment of prognostic factors in two studies. We rated all models as not usable in clinical practice (Table 2). Both risk of bias and applicability ratings were affected by inadequate reporting, which was a primary reason for “unclear” domain ratings.

Results of included studies
Heterogeneity of clinical characteristics, prognostic factors and methods, including the statistical approaches to multivariable modelling, precluded the statistical synthesis of the four development studies and limited the narrative synthesis of all five studies. Considering the studies' heterogeneity and poor performance against
the PROBAST criteria, we limited the presentation of data within our review to a table of key study characteristics without results (Table 1). For a more detailed table of the characteristics, including results, see the appendix (Table A.5).

The presented models differed greatly in various aspects including the number and composition of prognostic factors as well as in the presented statistics (see Table A.5). Only Hung et al.\textsuperscript{41} provided a prognostic index (Table A.5).

Conflicts of interest

Conflicts of interest were explicitly addressed only in two studies,\textsuperscript{40,43} which stated that there were none.

Discussion

Summary of main results

This systematic review includes five studies with a total of 387 patients that aimed to either develop\textsuperscript{39-41,43} or validate\textsuperscript{42} prognostic models for predicting outcomes in adults who undergo physiotherapy, with or without other conservative measures, for painful rotator cuff disorders.

The studies were heterogeneous in terms of the populations, the phases of research, the prognostic factors studied, the statistical approaches used and the results reported. These considerations ruled out meaningful statistical synthesis and imposed major limitations on narrative synthesis. Moreover, all of the studies were at high risk of bias and most raised “unclear” or “high” concerns about applicability. None of the models were ready for use in practice.
Overall completeness, applicability and usability of the evidence

The study populations were broadly relevant to the review question. Four studies investigated populations with impingement-related shoulder pain, implicitly including rotator cuff tears of varying completeness, except Hung et al, who excluded FTT. Merolla et al. exclusively studied rotator cuff tears, although it is unclear whether they incorporated PTT in this definition. However, applicability was compromised by unclear eligibility criteria in some studies, pertaining, for example, to frozen shoulder or rotator cuff tears. Also, in two studies the patient populations were selected, by dint of their agreement to participate in an RCT, which may have reduced external validity.

The physiotherapy intervention was insufficiently described to allow a judgement in Taheriazam et al. However, in the intervention group of Björnsson-Hallgren et al. and in the other three studies, the physiotherapy intervention was generally consistent with standard practice.

Less uniform was the selection of predictors, which was generally unjustified and diverse. In one case, prediction required measurement using specialised equipment (the FASTRAK motion analysis system) that would not be available in most clinical settings. Replicability and applicability of the models is likely to be reduced by the questionable clinimetric properties of some prognostic factor measurements, such as posterior shoulder tightness in Hung et al. and the application of arbitrary cut-points for categorizing continuous prognostic factors.

Some of our pre-specified outcomes were reported in some studies, including pain, shoulder disability and Global Perceived Change. Björnsson-Hallgren
reported the decision to undergo surgery. The remaining outcomes of interest for this review, including adverse events, HRQoL, sick leave and structural progression of tears, were either not reported or, in one case,\textsuperscript{42} reported too unclearly for extraction.

None of the four development studies\textsuperscript{39-41,43} reported any form of internal model validation; and none of these four was followed by an external validation, even though five and 10 years had elapsed since Hung et al.\textsuperscript{41} and Taheriazam et al.\textsuperscript{43} respectively. Lack of appropriate validation of prognostic models is a widely observed issue.\textsuperscript{47} There is good empirical evidence that models perform substantially less well in external, i.e. independent, samples, and that performance in external samples is more representative of clinical performance,\textsuperscript{28,48} so this presents a major obstacle to usability. The fifth study (Merolla et al.),\textsuperscript{42} though reportedly a validation, was seriously flawed in both concept and execution. Ultimately none of the studies has been assessed for clinical impact and, consequently, none of the models presented in the included studies is usable in clinical practice.

\textit{Quality of the evidence}

We evaluated risk of bias in five domains: participant selection, predictors, outcome, sample size and flow, and analysis. Our judgment of risk of bias was affected by a number of methodological issues (see results). Most of the identified deficiencies have been addressed extensively in the literature; several, including in particular those relating to the number of prognostic factors in relation to the number of outcome events and use of univariable analyses to select prognostic factors have been shown to result in invalid and unreliable models.\textsuperscript{49} Similarly, the use of statistical methods such as stepwise regression to select factors within the multivariable analysis has been criticized.\textsuperscript{49,50} This suggests that the presented
models are highly unlikely to produce valid and reliable predictions. Moreover, deficiencies such as unclear handling of missing data and the failure to consider overfitting of data, complexities in the data, evaluation of performance measures or non-linear relationships seriously hamper the judgement of the quality of the data and the models’ performance. The single “validation” study, by Merolla et al., was at high risk of bias in most domains.

An issue warranting special emphasis is the inclusion of prognostic factors in the outcome definition, i.e. the problem of incorporation bias through mathematical coupling, as this represents a conflict between risk of bias and applicability. The literature on incorporation bias primarily relates to diagnostic research. In that context, it relates to the interaction between index and reference tests. Mathematical coupling, which inherently occurs “when one variable directly or indirectly contains the whole or part of another” may either erroneously purport a relationship between the prognostic factor(s) and the outcome, or overestimate an existing relationship, thus inflating estimates of predictive performance. The conflict with applicability arises specifically because baseline and endpoint evaluation of a given outcome measure is standard clinical practice. This particularly applies to the increased use of PROMs in clinical practice and research. Moreover, in the present context, PROMs are among very few prognostic factors that have a basis in evidence In our review, this conflict was encountered in two studies, Kromer et al. and Taheriazam et al., which were both downgraded for risk of bias in the outcome domain. The described problem may be accommodated in the study design (e.g. by including a no-treatment control group as a point of reference) or addressed at the analysis stage, but should not be overlooked.
Potential biases in the review process

We sought to minimise bias in the review process by developing an *a priori* protocol that was registered with PROSPERO. In addition, the full protocol was lodged, *a priori*, with the Chair of the Research Governance and Ethics Committee of the School of Health and Social Care at Teesside University. We recorded any deviations from the protocol (Table A.1).

Our searches were comprehensive, and included several supplementary sources as well as the thorough inspection of all search results. The known difficulty of identifying prognosis research\textsuperscript{34,53} is reflected by the < 0.1\% yield of included studies from our initial results (Figure 1). Problems include the lack of appropriate indexing functions in the electronic databases and of current validated search filters. We identified a number of search filters for prognosis research,\textsuperscript{e.g.34,55,56} but had concerns about the currency of all but one,\textsuperscript{34} for Medline, that was purposely designed to identify prognostic *model* studies for systematic reviews. Applying this filter (amended by “prognos*”) significantly decreased the number of results in Medline, but nonetheless, in contrast to all other databases searched, retrieved all five studies that were included in this review. This suggests that this filter performs well.

Identification of relevant studies was also hampered by uninformative titles and abstracts, and inconsistent terminology compounded these difficulties, as has been noted by others.\textsuperscript{27,28} Although we restricted inclusion to reports in English, we did not impose a language restriction on our searches, and did not identify any non-English but clearly relevant studies.

Systematic reviewing of prognostic modelling studies is an evolving field, and the methodology is a work in progress. Nonetheless, in evaluating the studies we
referred to the latest recommendations of the PROGRESS partnership\textsuperscript{25} and, after piloting earlier versions, evaluated risk of bias and applicability using a near-definitive but unpublished version of PROBAST (R. Wolff, personal communication). The use of PROBAST was especially appealing to us because it is the first tool to specifically address risk of bias and applicability in prognostic model studies.

Agreements and disagreements with other studies or reviews

To our knowledge, this is the first systematic review to synthesize the evidence on primary prognostic model research in adults with rotator cuff disorders who are undergoing conservative treatment with physiotherapy. We identified two other prognostic systematic reviews addressing shoulder pain,\textsuperscript{37,38} but both aimed to synthesize evidence on individual prognostic factors rather than on prognostic models, and have minimal overlap with our own review, which has a single study\textsuperscript{41} in common with Chester et al.\textsuperscript{37} and none with Kuijpers et al.\textsuperscript{38} Of the two reviews, Chester et al.,\textsuperscript{37} like us, limited inclusion to studies investigating response to conservative treatment with physiotherapy, while Kuijpers et al.\textsuperscript{38} studied overall prognosis. Both reviews addressed shoulder pain in general and did not provide any subgroup analyses to allow for inferences about rotator cuff disorders. Thus, while evidence was found supporting a limited number of emerging factors including symptom duration, baseline function or disability,\textsuperscript{37,38} pain and age,\textsuperscript{38} the transferability of these findings to the population of interest in our review is unclear.

Conclusions

Implications for practice
There is no prognostic model ready to inform clinical practice on the prognosis of outcomes in adults who undergo physiotherapy, with or without other conservative measures, for painful rotator cuff disorders.

Implications for research

The complexity of prognostic modelling demands high levels of methodological expertise and clinical judgement, but particularly calls for the involvement, from the outset, of a statistician with expertise in the field. The composition of primary (but also secondary) research teams should therefore reflect this. Researchers should be receptive to developing methodologies which may improve the validity and reliability of prognostic models. Crucially, more attention should be paid to model validation, and ultimately, to the assessment of clinical impact.

The PROBAST tool, once publicly available, should facilitate the assessment of risk of bias and applicability in future systematic reviews of prognostic model studies. Further, both methods and reporting will benefit from adherence to the recommendations set out in the recent TRIPOD (Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis) statement. Further guidance for systematic reviews of prognostic model studies is now available through the CHARMS (Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies) checklist.

Acknowledgements

We gratefully acknowledge the assistance of Robert Wolff, Kleijnen Systematic Reviews Ltd., York, UK. Dr Wolff shared with us the pre-published versions of PROBAST which we piloted in the development of this review, and provided support
on their use. We also acknowledge the support from Iain Baird and Julie Hogg, Library and Information Services, Teesside University, Middlesbrough, UK, on the development of the search strategy; and Hanna Björnsson Hallgren for providing unpublished analysis data.

Disclosure of potential conflicts of interest:

All authors are contributors to an unfunded primary prognostic modelling study of rotator cuff disease treated by physiotherapy (registration nr. DRKS00004462). The systematic review was also unfunded.
References


List of tables and figures

List of tables

Table 1: Characteristics of included studies
Table 2: PROBAST (risk of bias + applicability) ratings

Tables as supplementary materials:
Table A.1: Deviations from protocol
Table A.2: Medline search strategy
Table A.3: Excluded studies
Table A.4: Ongoing studies
Table A.5: Characteristics and results of included studies – detailed table version

List of figures

Figure 1: Study flowchart
Table 1. Characteristics of included studies (alphabetical order)

<table>
<thead>
<tr>
<th>Study</th>
<th>DESIGN</th>
<th>SETTING</th>
<th>STARTPOINT</th>
<th>PARTICIPANTS</th>
<th>INTERVENTION</th>
<th>PROGNOSTIC FACTORS CONSIDERED</th>
<th>OUTCOME</th>
<th>ENDPOINT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Björnsson Hallgren 2014</strong></td>
<td>Cohort study derived from 2 group RCT; model development</td>
<td>Sweden; presumably outpatient</td>
<td>Not precisely defined; recruitment was from the waiting list for arthroscopic subacromial decompression (duration of symptoms ≥ 6 months)</td>
<td>N = 102 (data on 95); “subacromial pain”; mixed population: non-tear (69%), partial tear (22%), full tear (9%)</td>
<td>Exercise-based physiotherapy for both groups (specific versus control exercises) following initial steroid injection (duration 12 weeks)</td>
<td>N = 8: Constant score (quartiles), proximal humeral migration (yes/no), radiological determination of osteoarthritis [in the shoulder complex] (yes/no), cuff status (“intact”, “partial tear”, or “full tear”), subacromial calcification (yes/no), subacromial degeneration (yes/no), sex, treatment group (control versus specific)</td>
<td>Choice of surgery (yes/no, based on record of treatment) OUTCOME EVENTS n = 41</td>
<td>After 1 year (after inclusion or after surgery)</td>
</tr>
<tr>
<td><strong>Hung 2010</strong></td>
<td>Cohort (single-group); developmental; model development</td>
<td>Taiwan; presumably outpatient</td>
<td>Recruitment by an orthopaedics clinic or by “general announcements in the local internet media”; no further information was provided</td>
<td>N = 33 (of interest for the present review was a subgroup of 23 participants who showed “improvement”); “subacromial impingement syndrome”; presumably mixed population (rotator cuff tears were not excluded, but no further information was provided)</td>
<td>Standardized physical therapy programme (duration 6 weeks)</td>
<td>N = unclear; up to 60 may have been assessed covering the following predictors or categories: scapular kinematics, passive shoulder ROM, isometric strength, thoracic spine posture, posterior shoulder tightness, functional disability, symptom duration, compliance with treatment, age, height, weight</td>
<td>“Improvement” on 15-point GRCS, with dichotomisation into “improved” or “not improved”. OUTCOME EVENTS n = 23</td>
<td>After 6 weeks (conclusion of physical therapy treatment)</td>
</tr>
</tbody>
</table>

NOTES
*Unpublished analysis data specifies up to 97 observations

STATISTICAL ANALYSIS
Logistic regression
### Kromer 2014

**NOTES**  
† All potential prognostic factors were dichotomised; though the method of dichotomisation was pre-specified, its implementation was data-driven.

**DESIGN**  
Cohort study derived from 2 group RCT; model development

**SETTING**  
Germany; outpatient

**STARTPOINT**  
Presentation to a physiotherapist following referral by general practitioner or orthopaedic surgeon (duration of symptoms ≥ 4 weeks)

**PARTICIPANTS**  
90 (data for 88) “subacromial shoulder pain”; presumably tendinopathies & partial tears

**INTERVENTION**  
Both treatment groups included supervised exercises; the intervention group received additional treatment with manual mobilisations, individualised education & instruction on ADL (duration overall 12 weeks)

**PROGNOSTIC FACTORS CONSIDERED**  
N ≥ 7‡: Age, 11-point VNRS, FABQ-PA, PCS, Sex, SPADI-F, symptom duration

**OUTCOME**  
SPADI-F change score  
**OUTCOME EVENTS n = 88**  
**ENDPOINT**  
After 12 weeks (conclusion of intervention)

**SELECTION OF FACTORS FOR MULTIVARIABLE MODELLING**  
It is unclear what predictors were initially considered. Multicollinearity was assessed among the seven predictors that are specified in the report (cut-off r >/= .5); in case of a correlation, the “most easily obtainable variable in clinical practice” was chosen for further analysis; selection was done irrespective of the statistical significance of univariable correlations of predictors with the outcome.  
**WITHIN MULTIVARIABLE MODELLING** backward regression

**STATISTICAL ANALYSIS**  
Linear regression (hierarchical)

**NOTES**  
‡ The narrative implies that there were other, unspecified, predictors.

### Merolla 2011§

**DESIGN**  
Cohort (single-group); model validation

**SETTING**  
Italy; outpatient

**STARTPOINT**  
Diagnosis of a symptomatic rotator cuff tear by a shoulder surgeon

**PARTICIPANTS**  
N = 60 (of interest for the present review was a subgroup of 33 participants who were treated conservatively); “symptomatic rotator cuff tears” (presumably both partial & full-thickness)

**INTERVENTION**  
Treatment included pain control, passive mobilisation, supervised exercises and laser therapy (overall duration unclear)

**PROGNOSTIC FACTORS CONSIDERED**  
N ≥ 17. Acromiohumeral interval (>/> 7mm), active ROM (>/> 90°, though the movements to which this applied were unspecified), age (>/> 60 years), bilateral tear (yes or no), drop sign (yes or no), long head of biceps status (“normal”, “rupture”, “instability”), overhead sport (yes or no), previous rehabilitation (yes or no), scapular dyskinesis (yes or no), shoulder trauma (</> 6 months), subscapularis tear (yes or no), type of tear (“complete”, “partial”), working activity (“light”, “heavy”), working compensation (yes or no), Passive stiffness, measured goniometrically (“none or mild”, “moderate”, “severe”), rotator cuff fatty infiltration (Grades 0-I, II or III), & rotator cuff muscle atrophy (Grades I, II, III or IV)

**OUTCOMES**  
Constant score, “subjective satisfaction” (by a “nominal” scale), pain (by VAS). It is unclear whether all were used for the validation of the model. ‘Election of surgery’ & QoL also appear to have been assessed, but were not pre-specified
<table>
<thead>
<tr>
<th>Outcomes in the Methods. OUTCOME EVENTS 33 for continuous outcomes (conservatively treated participants)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ENDPOINT</strong></td>
</tr>
<tr>
<td><strong>STATISTICAL ANALYSIS</strong></td>
</tr>
<tr>
<td><strong>NOTES</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Taheriazam 2005</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DESIGN</strong></td>
</tr>
<tr>
<td><strong>SETTING</strong></td>
</tr>
<tr>
<td><strong>STARTPOINT</strong></td>
</tr>
<tr>
<td><strong>PARTICIPANTS</strong></td>
</tr>
<tr>
<td><strong>INTERVENTION</strong></td>
</tr>
<tr>
<td><strong>PROGNOSTIC FACTORS CONSIDERED</strong></td>
</tr>
<tr>
<td><strong>OUTCOMES</strong></td>
</tr>
<tr>
<td><strong>ENDPOINT</strong></td>
</tr>
<tr>
<td><strong>SELECTION OF FACTORS FOR MULTIVARIABLE MODELLING</strong></td>
</tr>
<tr>
<td><strong>WITHIN MULTIVARIABLE MODELLING</strong></td>
</tr>
<tr>
<td><strong>STATISTICAL ANALYSIS</strong></td>
</tr>
<tr>
<td><strong>NOTES</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ABBREVIATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADL = Activities of Daily Living, FABQ-PA = Fear Avoidance Beliefs Questionnaire Physical Activity subscale, FLEX-SF = Flexilevel Scale of Shoulder Function, GRCs = Global Rating of Change Scale, NI = No information, NSAIDs = Non-Steroidal Anti-Inflammatory Drugs, QoL = Quality of Life, RCT = Randomised Controlled Trial, ROM = Range of Motion, SD = Standard Deviation, SLAP = Superior Labral Anterior to Posterior, SPADI-F = Shoulder Pain &amp; Disability Index Function subscale, VAS = Visual Analogue Scale, VNRS = Visual Numeric Rating Scale.</td>
</tr>
</tbody>
</table>
Table 2: PROBAST (risk of bias and applicability) ratings

<table>
<thead>
<tr>
<th>Study</th>
<th>RISK OF BIAS</th>
<th>APPLICABILITY CONCERNS</th>
<th>OVERALL JUDGEMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Björnsson Hallgren 2014</td>
<td>☑</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td>Kromer 2014</td>
<td>☑</td>
<td>☑</td>
<td>☑</td>
</tr>
</tbody>
</table>

☑ Low risk/concerns ☐ High risk/concerns ☞ Unclear risk/concerns
Table A.1: Deviations from protocol

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Difference with justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome</td>
<td>We added “need for surgery” to the outcomes upon noticing that it was used in a number of potentially relevant studies; and thus obviously viewed as an outcome of interest to other researchers in this field.</td>
</tr>
<tr>
<td>Presentation of prognostic model</td>
<td>Upon finding that incomplete reporting was a major issue, we added as a requirement for inclusion that the final prognostic model (or the most complete model including main effects for all prognostic factors) was either fully reported or that a full report was provided on request.</td>
</tr>
<tr>
<td>Author contact</td>
<td>We planned to contact study authors for unreported study details and data, but later decided to limit author contact to the clarification of issues related to study eligibility (at the second screening step), because we considered it very unlikely that obtaining the missing data would make any important differences to the outcome and conclusions of our review.</td>
</tr>
</tbody>
</table>
Table A.2: Medline search strategy (EBSCO format)

<table>
<thead>
<tr>
<th>Step</th>
<th>Search Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>(MH “Shoulder” OR MH “Shoulder Pain” OR shoulder) AND (MH Tendinopathy OR (“soft tissue” OR tendon* OR tendin* OR imping* OR rotator OR cuff).ti,ab)) OR (supraspinatus OR infraspinatus OR “teres minor” OR subscapularis OR “rotator cuff” OR subacromial*).ti,ab OR MH “Shoulder Impingement Syndrome” OR MH “Rotator Cuff”</td>
</tr>
<tr>
<td>S2</td>
<td>MH “Physical Therapy Modalities+” OR MH “Rehabilitation+” OR (“physical therap*” OR physiotherap* OR exercis* OR “manual therap*” OR ”manipulative therap*” OR mobilis* OR rehab* OR conservative* OR non-operat* OR nonoperat* OR non-surg* OR nonsurg*).ti,ab</td>
</tr>
<tr>
<td>S3</td>
<td>validat* OR TI predict*.ti OR rule* OR (predict* AND (outcome* OR risk* OR model*)) OR ((history OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor*) AND (predict* OR model* OR decision* OR identif* OR prognos*)) OR (decision* AND (model* OR clinical* OR MH “Logistic Models”)) OR (prognostic AND (history OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor* OR model*))</td>
</tr>
<tr>
<td>S4</td>
<td>stratification OR MH ”ROC Curve” OR discrimination OR discriminate OR c-statistic OR ”c statistic” OR area under the curve OR AUC OR calibration OR indices OR algorithm OR multivariable</td>
</tr>
<tr>
<td>S5†</td>
<td>prognos*.ti,ab</td>
</tr>
<tr>
<td>S6</td>
<td>S1 AND S2 AND (S3 OR S4 OR S5)</td>
</tr>
</tbody>
</table>

† Prognosis research filter as proposed by Geersing et al. (2012; see review reference list) (clinical prediction model studies, Ingui filter OR update (S3 OR S4))

† Amendment to the Geersing search filter (S3 OR S4 OR S5)
<table>
<thead>
<tr>
<th>Nr</th>
<th>Study</th>
<th>Main reasons for exclusion – criterion categories* (in brackets) and explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Audenaert A, de Mey E, Reniers G. Patient variables determining treatment protocol and related economical impact in occupational rotator cuff tears. WSEAS Trans Biol Biomed 2012;9:24–33.</td>
<td>(Po) Traumatic population (all participants had experienced a &quot;posttraumatic rotator cuff tear in an industrial accident&quot;)</td>
</tr>
<tr>
<td>2</td>
<td>Bartolozzi A, Andreychik D, Ahmad S. Determinants of outcome in the treatment of rotator cuff disease. Clin Orthop Relat Res 1994:90–7.</td>
<td>(I) It is not made explicit that all participants received physical therapy, and as “the three treatment options (PT, injection and NSAIDs) were also assessed” as predictive factors, it seems unlikely (S/A) Retrospective study</td>
</tr>
<tr>
<td>3</td>
<td>Bokor DJ, Hawkins RJ, Huckell GH, Angelo RL, Schickendantz MS. Results of nonoperative management of full-thickness tears of the rotator cuff. Clin Orthop Relat Res 1993:103–10.</td>
<td>(Po) Only 24% were atraumatic (S/A) No multivariable prognostic modelling; retrospective</td>
</tr>
<tr>
<td>6</td>
<td>Chard MD, Sattelle LM, Hazleman BL. The long-term outcome of rotator cuff tendinitis--a review study. Br J Rheumatol 1988;27:385–9.</td>
<td>(Po) Only 21% were atraumatic. (I) Only 16% of the sample underwent physiotherapy and there is no separate analysis for this subgroup. (S/A) No multivariable prognostic modelling</td>
</tr>
<tr>
<td>9</td>
<td>Curry EJ, Matzkin EE, Dong Y, Higgins LD, Katz JN, Jain NB. Structural Characteristics Are Not Associated With Pain and Function in Rotator Cuff Tears: The ROW Cohort Study. Orthop J Sport Med 2015;3.</td>
<td>(S/A) Cross-sectional study; no multivariable prognostic modelling</td>
</tr>
<tr>
<td>Number</td>
<td>Reference</td>
<td>Additional Information</td>
</tr>
<tr>
<td>--------</td>
<td>---------------------------------------------------------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>10</td>
<td>Deutscher D, Horn SD, Dickstein R, Hart DL, Smout RJ, Gutvirtz M, et al.</td>
<td>(Po) Population not condition-specific (shoulder pain as one out of four musculoskeletal impairment groups categories)</td>
</tr>
<tr>
<td>11</td>
<td>Ekeberg OM, Bautz-Holter E, Juel NG, Engbretnsen K, Kvalheim S, Brox JI.</td>
<td>(I) No defined physiotherapy treatment: was allowed if started, but was not followed. Only some patients had physiotherapy (see primary RCT report)</td>
</tr>
<tr>
<td>14</td>
<td>Gagnier JJ, Robbins C, Carpenter JE, Bedi A, Miller B. A Prospective Cohort Study of Patients Treated Surgically or Non-Surgically for Full-thickness Rotator Cuff Tears. Orthop J Sport Med 2014;2 suppl. doi:10.1177/232596714S00059.</td>
<td>(TP) Extended abstract; no published full study report available (may be linked with Kweon et al. 2015)</td>
</tr>
<tr>
<td>16</td>
<td>Goldberg BA, Nowinski RJ, Matsen FA. Outcome of nonoperative management of full-thickness rotator cuff tears. Clin Orthop Relat Res 2001;(382):99–107.</td>
<td>(I) No mention of supervised physiotherapy or of any involvement of physiotherapists (treatment consisted of a home exercise program only) (S/A) There is a paragraph relating to prediction, but it is completely unclear how these results were derived. There is no reporting of multivariable modelling, and no mention of such in the methods</td>
</tr>
<tr>
<td>18</td>
<td><strong>Hawkins</strong> RH, Dunlop R. Nonoperative treatment of rotator cuff tears. Clin Orthop Relat Res 1995:178–88.</td>
<td>(Po) 64% of cases were traumatic (O) The outcome variable is patient satisfaction, which is not an outcome of interest in this review</td>
</tr>
<tr>
<td>25</td>
<td><strong>Kulenkampff</strong> H-A, Reichelt A. Clinical course of ruptures of the rotator cuff after conservative therapy. Orthopadische Prax 1990;26:493–6.</td>
<td>(L) Full text in German (S/A) Not a prognostic model study</td>
</tr>
<tr>
<td>26</td>
<td>Kweon C, Gagnier JJ, Robbins CB, Bedi a., Carpenter JE, Miller BS. Surgical Versus Nonsurgical Management of Rotator Cuff Tears: Predictors of Treatment Allocation. Am J Sports Med 2015;8–13. doi:10.1177/0363546515593954.</td>
<td>(S/A) Not designed to follow a course of conservative treatment with physiotherapy over a defined period of time (allocation to surgery could have happened any time); although part of a prospective cohort study, the prognostic assessment seems like a case control comparison.</td>
</tr>
<tr>
<td>28</td>
<td><strong>Maman E, Harris C, White L, Tomlinson G, Shashank M, Boynton E.</strong> Outcome of nonoperative treatment of symptomatic rotator cuff tears monitored by magnetic resonance imaging. J Bone Joint Surg Am 2009;91:1898–906. doi:10.2106/JBJS.G.01335.</td>
<td>(Pr), (S/A): No multivariable prognostic modelling related to the variables of interest for this review: the relationship between baseline variables and changes in tear size was evaluated by simple percentage comparisons. Logistic regression was only used to assess the relationship between progression in tear size and elapsed time between a participant’s first and final MRI scan; retrospective study</td>
</tr>
<tr>
<td>29</td>
<td><strong>McCreesh K.</strong> Evidence based prognosis setting in the case of a conservatively managed rotator cuff tear. Physiother Irel 2007;28:31–5.</td>
<td>(S/A) Case study</td>
</tr>
<tr>
<td>31</td>
<td><strong>Morag Y, Jamadar DA, Miller B, Brandon C, Gandikota G, Jacobson JA.</strong> Morphology of large rotator cuff tears and of the rotator cable and long-term shoulder disability in conservatively treated elderly patients. J Comput Assist Tomogr 2013;37:631–8.</td>
<td>(I) 80% of the sample received physiotherapy, but these are not separately reported (S/A) No multivariable prognostic modelling</td>
</tr>
<tr>
<td>33</td>
<td><strong>Notarnicola A, Maccagnano G, Tafuri S, Fiore A, Margiotta C, Pesce V, et al.</strong> Prognostic factors of extracorporeal shock wave therapy for tendinopathies. Musculoskelet Surg 2015. doi:10.1007/s12306-015-0375-y.</td>
<td>(Po) Mixed population of various musculoskeletal tendon complaints including rotator cuff tendinitis, combined analysis (no difference in response to treatment was found related to the different tendons) (I) Following a course of extracorporeal shockwave therapy; physiotherapy treatment was documented, but was not standard element of treatment</td>
</tr>
<tr>
<td>34</td>
<td><strong>Ottaviani M, Mele G.</strong> Epidemiological, clinical and diagnostic study of rotator cuff rupture. Riabilitazione 1998;31:17–24.</td>
<td>(L) Full text in Italian (S/A) Presumably anyway not a multivariable prognostic modelling study</td>
</tr>
<tr>
<td>35</td>
<td><strong>Rahme H, Solem-Bertoft E, Westerberg CE, Lundberg E, Sörensen S, Hilding S.</strong> The subacromial impingement syndrome. A study of results of treatment with special emphasis on predictive factors and pain-generating mechanisms. Scand J Rehab Med;30:253–62.</td>
<td>(Po) 24% of overall sample were post-trauma (subgroup data not reported) (O) The outcome is the success of surgery (i.e. only surgically treated patients were evaluated by multivariable regression analysis)</td>
</tr>
<tr>
<td>36</td>
<td><strong>Rowe CR.</strong> Ruptures of the rotator cuff: selection of cases for conservative treatment.</td>
<td>(S/A) Not a prognostic modelling study</td>
</tr>
<tr>
<td></td>
<td>Reference</td>
<td>Notes</td>
</tr>
<tr>
<td>---</td>
<td>---------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>38</td>
<td>Safran O, Schroeder J, Bloom R, Weil Y, Milgrom C. Natural history of nonoperatively treated symptomatic rotator cuff tears in patients 60 years old or younger. Am J Sports Med 2011;39:710–4. doi:10.1177/0363546510393944.</td>
<td>(Po) 53% were post-traumatic (I) No mention of physiotherapy; not following a defined course of conservative treatment with physiotherapy (&quot;natural progression&quot;) (S/A) There appears to be no prognostic modelling</td>
</tr>
<tr>
<td>39</td>
<td>Samilson RL, Binder WF. Symptomatic full thickness tears of rotator cuff. An analysis of 292 shoulders in 276 patients. Orthop Clin North Am 1975;6:449–66.</td>
<td>(Po) 82% were post-traumatic (I) An unspecified proportion received physiotherapy and there is no discrete physiotherapy subgroup (S/A) Not a prognostic modelling study</td>
</tr>
<tr>
<td>43</td>
<td>Solomon DH, Bates DW, Schaffer JL, Horsky J, Burdick E, Katz JN. Referrals for musculoskeletal disorders: patterns, predictors, and outcomes. J Rheumatol 2001;28:2090–5.</td>
<td>(I) Treatment unspecified (not all patients received physiotherapy); i.e. not following a defined course of conservative treatment with physiotherapy (O) Outcome of interest (&quot;referral&quot; to a secondary care specialist) not of interest for this review</td>
</tr>
<tr>
<td>45</td>
<td>Van Der Windt DAWM, Koes BW, Boeke AJP, Devillé W, De Jong B a, Bouter LM. Shoulder</td>
<td>(I) Not following a defined course of conservative treatment with physiotherapy:</td>
</tr>
<tr>
<td></td>
<td>disorders in general practice: Prognostic indicators of outcome. Br J Gen Pract 1996;46:519–23.</td>
<td>not all patients (in the rotator cuff tendinitis group) had physiotherapy</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

*Criterion categories:
Po = population, I = Intervention(s), O = Outcome(s), S/A = Study design/Analysis, Pr = prognostic factors, L = Language; TP = Type of publication
### Table A.4: Ongoing studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ICTRP (study ID and title), ordered by ID</strong></td>
<td></td>
</tr>
<tr>
<td>ACTRN12615000351516</td>
<td>Pain modulation characteristics in people with shoulder impingement and predictors of successful outcomes following physiotherapy treatment. ITRP. Available at: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=ACTRN12615000351516">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=ACTRN12615000351516</a> [last accessed 22 Oct 2015]</td>
</tr>
<tr>
<td>DRKS00004462</td>
<td>Predicting the outcome of conservative treatment with physiotherapy for shoulder pain in the presence of atraumatic partial-thickness tears of the rotator cuff. ITRP. Available at: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00004462">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00004462</a> [last accessed 22 Oct 2015]</td>
</tr>
<tr>
<td>NCT00762580</td>
<td>Features to Predict Success With Nonoperative Treatment of Patients With Rotator Cuff Tears (MOON). ITRP. Available at: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00762580">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00762580</a> [last accessed 22 Oct 2015]</td>
</tr>
<tr>
<td>NCT02510352</td>
<td>Cohort of Patients With a Symptomatic Rotator Cuff Tear Treated Without Surgical Repair. ITRP. Available at: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02510352">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02510352</a> [last accessed 22 Oct 2015]</td>
</tr>
<tr>
<td><strong>Published protocol (first author (year))</strong></td>
<td></td>
</tr>
</tbody>
</table>
Table A.5 Characteristics and results of included studies – detailed version

<table>
<thead>
<tr>
<th><strong>Björnsson Hallgren 2014</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OBJECTIVE</strong></td>
</tr>
<tr>
<td><strong>DESIGN</strong></td>
</tr>
<tr>
<td><strong>PHASE OF RESEARCH</strong></td>
</tr>
<tr>
<td><strong>SETTING</strong></td>
</tr>
<tr>
<td><strong>STARTPOINT</strong></td>
</tr>
<tr>
<td><strong>PARTICIPANTS</strong></td>
</tr>
<tr>
<td><strong>TYPE OF DISORDER</strong></td>
</tr>
<tr>
<td><strong>INTERVENTION</strong></td>
</tr>
<tr>
<td><strong>PROGNOSTIC FACTORS CONSIDERED</strong></td>
</tr>
<tr>
<td><strong>OUTCOME</strong></td>
</tr>
<tr>
<td><strong>ENDPOINT</strong></td>
</tr>
<tr>
<td><strong>SELECTION OF FACTORS FOR MULTIVARIABLE MODELLING</strong></td>
</tr>
<tr>
<td><strong>WITHIN MULTIVARIABLE MODELLING</strong></td>
</tr>
<tr>
<td><strong>STATISTICAL ANALYSIS</strong></td>
</tr>
<tr>
<td><strong>MOST COMPLETE MODEL INCLUDING MAIN EFFECTS FOR ALL PROGNOSTIC MODELS</strong></td>
</tr>
<tr>
<td><strong>Pseudo R²</strong>: 0.28</td>
</tr>
<tr>
<td><strong>Predictor/statistics‡</strong></td>
</tr>
<tr>
<td>Intact cuff</td>
</tr>
<tr>
<td>PTT</td>
</tr>
<tr>
<td>FTT</td>
</tr>
<tr>
<td>Control vs. specific</td>
</tr>
<tr>
<td>CM 1. quartile</td>
</tr>
<tr>
<td>CM 2. quartile</td>
</tr>
<tr>
<td>CM 3. quartile</td>
</tr>
<tr>
<td>CM 4. quartile</td>
</tr>
<tr>
<td>Calcification</td>
</tr>
<tr>
<td>Degeneration</td>
</tr>
<tr>
<td>Women</td>
</tr>
<tr>
<td>Regression constant</td>
</tr>
<tr>
<td><strong>FURTHER EVALUATION OF MODEL PERFORMANCE</strong></td>
</tr>
<tr>
<td><strong>PROGNOSTIC INDEX/STATEMENT</strong></td>
</tr>
</tbody>
</table>
**STUDY AUTHORS’ CONCLUSIONS**

“The severity of shoulder disability at baseline and the presence of a full-thickness tear seem to influence outcome and the need for surgery.”

**NOTES**

- * Unpublished analysis data specifies up to 97 observations
- † Adjustment variables
- ‡ Based on unpublished analysis data
- § Model based on 93 observations; values rounded to two decimal places

---

**OBJECTIVE**

“...to identify the shoulder kinematic and impairment of the patients who are more likely to respond to physical therapy;”

**DESIGN**

Cohort (single-group); developmental; presumably consecutive recruitment (no information provided)

**SETTING**

Taiwan; orthopaedic [presumably outpatient] clinic in a national university hospital

**PHASE OF RESEARCH**

development

**STARTPOINT**

Recruitment by an orthopaedics clinic or by “general announcements in the local internet media”; no further information provided

**PARTICIPANTS**

N = 33 (of interest for the present review was a subgroup of 23 participants who showed “improvement”) **TYPE OF DISORDER**

“subacromial impingement syndrome”; presumably mixed population (rotator cuff tears were not excluded, but no further information was provided) **MEAN AGE** 23.3 years; **SEX** 100% male

**INTERVENTION**

Standardised physical therapy programme **DURATION** 6 weeks

**PROGNOSTIC FACTORS CONSIDERED**

N unclear; up to 60 may have been assessed covering the following predictors or categories: scapular kinematics, passive shoulder ROM, isometric strength, thoracic spine posture, posterior shoulder tightness, functional disability, symptom duration, compliance with treatment, age, height, weight

**OUTCOME**

“Improvement” on 15-point GRCS from -7 (“a very great deal worse”) to +7 (“a very great deal better”), with dichotomisation into “improved” (≥ +4) or “not improved” (≤ +3).

**ENDPOINT**

After 6 weeks (conclusion of physical therapy treatment)

**SELECTION OF FACTORS FOR MULTIVARIABLE MODELLING**

“Variables from the shoulder kinematics and clinical impairments were tested for their relationship with the reference outcome using independent sample t-tests. Variables with a significant level of p < 0.10 may be retained as potential predictor variables.” **WITHIN MULTIVARIABLE MODELLING**

Stepwise regression.

**STATISTICAL ANALYSIS**

Logistic regression. Apparently, two models were calculated.

**FINAL MODEL**

N outcome events: 23

Nagelkerke R²: 0.73

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cut-off</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLEX-SF score</td>
<td>&lt; 41&quot;</td>
</tr>
<tr>
<td>Scapular internal rotation at 30° shoulder elevation (descending phase, unloaded)</td>
<td>&lt; 0.7&quot;</td>
</tr>
<tr>
<td>Serratus anterior force as % of body weight</td>
<td>&lt; 27%</td>
</tr>
</tbody>
</table>

**FURTHER EVALUATION OF MODEL PERFORMANCE** (including internal and external validation) Probability of improvement (%) was evaluated
for one, two or all of the factors in the final model: 1+: 69; 2+: 88; 3+: 100

**PROGNOSTIC INDEX/STATEMENT**

“...a subject with SAIS who meets 3 criteria (FLEX-SF score < 41, muscle force of serratus anterior < 27.4% body weight, degree of scapular internal rotation at 30° shoulder elevation < 0.7 degree) at baseline has a probability of 100% of demonstrating improvement at 6-week follow-up.”

**STUDY AUTHORS’ CONCLUSIONS**

See above

**NOTES**

²All potential prognostic factors were dichotomised using cut-points derived from ROC analyses. ³Apparently an adjustment variable. ⁴Resulting values from sensitivity and specificity ROC analysis.
FURTHER EVALUATION OF MODEL PERFORMANCE (including internal and external validation) None presented

PROGNOSTIC INDEX/STATEMENT None presented

STUDY AUTHORS’ CONCLUSIONS „In patients with SPS, fear-avoidance beliefs measured at baseline“ appear to be significantly associated with baseline disability but not with not with disability change scores after 3 months.“ „..., the regression model for the disability change score after 3 months clearly identified duration of complaints and baseline disability as the only significant variables.“

NOTES **The narrative implies that there were other, unspecified, predictors. ††Apparently an adjustment variable. ‡‡CIs contain inaccuracies (see italicized values)

Merolla 2011§§

OBJECTIVE „...to validate a prognostic score to predict which patients could have a good and stable outcomes with non operative treatment.“

DESIGN Cohort (single-group); consecutive recruitment PHASE OF RESEARCH validation

SETTING Italy; [outpatient clinic of] hospital department of shoulder & elbow surgery STUDY DATES unspecified

STARTPOINT Diagnosis of a symptomatic rotator cuff tear by a shoulder surgeon

PARTICIPANTS N 60 (of interest for the present review was a subgroup of 33 participants who were treated conservatively) TYPE OF DISORDER symptomatic rotator cuff tears (presumably both partial & full-thickness) MEAN AGE 52.6 years SEX 60% male

INTERVENTION Treatment was structured into different phases and included pain control, passive mobilisation, supervised exercises and laser therapy DURATION overall duration unclear

PROGNOSTIC FACTORS CONSIDERED N ≥ 17. There was no regression. Acromiohumeral interval (>/< 7mm), active ROM (>/< 90°, though the movements to which this applied were unspecified), age (>/< 60 years), bilateral tear (yes or no), drop sign (yes or no), long head of biceps status (“normal”, “rupture”, “instability”), overhead sport (yes or no), previous rehabilitation (yes or no), scapular dyskinesis (yes or no), shoulder trauma (</> 6 months), subscapularis tear (yes or no), type of tear (“complete”, “partial”), working activity (“light”, “heavy”), working compensation (yes or no), Passive stiffness, measured goniometrically (“none or mild”, “moderate”, “severe”), rotator cuff fatty infiltration (Grades 0-I, II or III), & rotator cuff muscle atrophy (Grades I, II, III or IV)

OUTCOMES Constant score, “subjective satisfaction” by a 0-100 “nominal” scale, & pain by VAS. It is unclear whether all were used for the validation of the model. ‘Election of surgery’ & QoL also appear to have been assessed, but were not pre-specified outcomes in the Methods.

ENDPOINT Unclear. Outcomes were measured at 6, 9 & 12 months, but the prognosis aspect may have been assessed at 12 months only.

STATISTICAL ANALYSIS “Student’s t-test was used to highlight significant differences between pre- and post-rehabilitation program scores.”

VALIDATION STATISTICS N outcome events: 33 for continuous outcomes (conservatively treated participants), unclear for categorised outcomes No validation statistics presented. Mean prediction score (SD) at follow-up:
CONSIDERATION OF CHANGES TO ORIGINAL MODEL

No information

STUDY AUTHORS’ CONCLUSIONS

«... the outcomes of our study support the assumption that a predictive prognostic score may guarantee a rational approach in the management of subjects with [cuff] tears, especially in elderly who continue to have the higher rate of recurrence and therefore could be well treated with standard conservative therapies.»

«Since the patients who benefit from conservative treatment had a score lower than 13 points, we identified this values as a “cut-off” score to predict a good results by conservative management of [cuff] tear.»

NOTES

Unclear & incomplete reporting seriously hindered data extraction.

Taheriazam 2005

OBJECTIVE

“...to determine the prognostic factors associated with the response to conservative therapy of subacromial impingement syndrome.”

DESIGN

Cohort (single-group); consecutive recruitment

PHASE OF RESEARCH development

SETTING

Iran outpatient orthopaedic clinic

STUDY DATES enrolment took place from March 2001 to February 2002

STARTPOINT

New diagnosis of impingement syndrome

PARTICIPANTS

N 102 (data for 89) TYPE OF DISORDER subacromial impingement syndrome (NI on whether or not rotator cuff tears were included)

MEAN AGE 56.4 years

SEX 51% male

INTERVENTION

Treatment was based on a standardised protocol including oral NSAIDs, up to two local steroid injections and a supervised physical therapy program; DURATION overall presumably 12 months

PROGNOSTIC FACTORS CONSIDERED

N 8 Acromial morphology (type I, II or III), acromial spur (present, absent), active ROM into flexion & abduction (implicitly measured goniometrically, but converted into ordinal data for analysis, as “normal”, “mildly impaired”, “moderately impaired”, or “severely impaired”), age, Constant score, dominant shoulder involvement (yes or no), sex, symptom duration.

OUTCOMES

Constant score

ENDPOINT

After 12 months (follow-up visit at clinic)

SELECTION OF FACTORS FOR MULTIVARIABLE MODELLING All eight predictors were included in the multivariable analysis, irrespective of the statistical significance of univariable correlations of predictors with the outcome. WITHIN MULTIVARIABLE MODELLING After the initial inclusion of all predictors, further modelling was based on the statistical significance of the regression coefficients. Among the three remaining predictors, three further multivariable models were then calculated.

STATISTICAL ANALYSIS

Linear regression, presumably four multivariable models were calculated.

FINAL MODEL

N outcome events: 89

R² adjusted: .68

Acromial morphology SEE = 0.76

Duration of symptoms

Baseline Constant score

Normal distribution of residuals was assessed (Kolmogorov-Smirnov
test): p = .3

**FURTHER EVALUATION OF MODEL PERFORMANCE** (including internal and external validation) None presented

<table>
<thead>
<tr>
<th>PROGNOSTIC INDEX/STATEMENT</th>
<th>None presented</th>
</tr>
</thead>
<tbody>
<tr>
<td>CALIBRATION, DISCRIMINATION, VALIDATION</td>
<td>None presented</td>
</tr>
<tr>
<td>PROGNOSTIC INDEX/STATEMENT</td>
<td>None presented</td>
</tr>
<tr>
<td>STUDY AUTHORS’ CONCLUSIONS</td>
<td>“We found that the predictive value of the pretreatment Constant score could be empowered by taking into account the effects of acromion morphology and pretreatment symptom duration. This is quantitatively shown by better fitness of the 3-variable model than the univariate models.”</td>
</tr>
<tr>
<td>NOTES</td>
<td>As reported by the authors, but there is a discrepancy. Of 128 eligible patients, 93 consented &amp; 13 were excluded from the analysis, giving a sample of 80. Erroneously analysed as continuous data in the regression.</td>
</tr>
</tbody>
</table>

**ABBREVIATIONS**
ADL = Activities of Daily Living, FABQ-PA = Fear Avoidance Beliefs Questionnaire Physical Activity subscale, FLEX-SF = Flexilevel Scale of Shoulder Function, GRCS = Global Rating of Change Scale, NI = No information, NSAIDs = Non-Steroidal Anti-Inflammatory Drugs, QoL = Quality of Life, RCT = Randomised Controlled Trial, ROC = Receiver Operating Characteristic, ROM = Range of Motion, SD = Standard Deviation, SEE = Standard Error of the Estimate, SLAP = Superior Labral Anterior to Posterior, SPADI-F = Shoulder Pain & Disability Index Function subscale, VAS = Visual Analogue Scale, VNRS = Visual Numeric Rating Scale.
Figure 1: Search and selection flow diagram (adopted from\textsuperscript{59})

Records identified through \textit{database searches}
cut-off 15 Nov 2014 (deduplicated initial
search + updates)
\(n = 5,889\)

Records identified from other sources
\(n = 12\)

Screening

Records screened
\(n = 5,901\)

Records excluded
\(n = 5,847\)

Ongoing studies (see Table A.4)
\(n = 8\)

Eligibility

Full-text articles assessed for eligibility
\(n = 54\)

Studies excluded
\(n = 49\)
(for reasons see table A.3)
Studies awaiting classification
\(n = 0\)

Included

Studies included in narrative synthesis
\(n = 5\)

Studies included in quantitative synthesis (meta-analysis)
\(n = 0\)