MANUSCRIPT FOR SUBMISSION TO *PHYSICAL THERAPY*

ACCEPTED VERSION (version 04 Nov 2015; accepted 22 Nov 2015)

Title

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

Systematic review registration number (PROSPERO)

CRD42014008973

Authors

1. Cordula Braun[1,2] , MSc
2. Nigel C. Hanchard[1], PhD
3. Alan M. Batterham[1], PhD
4. Helen H. Handoll[1], DPhil
5. Andreas Betthäuser[3], Dr med

Affiliations

[1]Health and Social Care Institute, School of Health and Social Care, Teesside University, Middlesbrough, United Kingdom;

[2] Department of Health (Physiotherapy), hochschule 21, Buxtehude, Germany;

[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.

Body of article

Title

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.

Data extraction

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.

Data synthesis

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

Limitations

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.

Conclusions

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.[1] 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.[2,3] Most shoulder complaints (29% to 85%) involve the subacromial-subdeltoid bursa and rotator cuff.[4,5] The pathology is diverse, reflecting a degenerative continuum from tendinopathy to partial (PTT) or full-thickness tears (FTT).[6] Rotator cuff tears, in particular, have a reported prevalence of over 40% in symptomatic shoulder pain populations[7] and are strongly correlated with age.[8] Clinical features of rotator cuff disorders may include pain, abnormalities on tests of rotator cuff function and integrity,[9] and significantly impaired shoulder function and health-related quality of life.[10,11] While diagnosis of rotator cuff disorders is based on clinical signs and symptoms,[9] verification of a rotator cuff tear requires diagnostic imaging (e.g. ultrasonography, magnetic resonance imaging).[12]

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.[13-15] Direct comparisons of conservative versus surgical treatment[16-19] have not shown clinically relevant differences between groups. Nonetheless, the rates of surgical intervention for rotator cuff disease have considerably increased in many countries.[20,21] 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”.[22]

The importance of predicting which patients will respond to particular treatments is increasingly recognised and has stimulated interest in prognosis and prognosis research.[23] There has been a corresponding development in prognosis research methodology.[24-26] Prognosis research aims to predict clinical outcomes in individual patients.[25] 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.[27,28] 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: NPI = (0.2 x tumor diameter (cm)) + lymph node stage + tumour grade.[29] 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).[28,30] External validation is a crucial step before a model can be considered usable in clinical practice.[28]

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[25] and, complementarily, the Cochrane Prognosis Methods Group.[26] We used PROGRESS terminology where possible.[27] This review is based on an a priori protocol, registered in PROSPERO, the International Prospective Register of Systematic Reviews[31] (registration nr. CRD42014008973), and available at www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014008973#.VTodbmYom1k. 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.

Participants

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,[32] 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,[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;[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),[35] 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.[36] 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,[37,38] five from our previous searches, and one from personal

258 communication.[39] We included five studies[39-43] and excluded 49 (see Table A.3 for
259 further details). The most frequent reason for exclusion was a lack of multivariable
260 prognostic modelling. We identified (by protocol or registry entries) eight clearly or
261 potentially relevant ongoing studies (see Figure 1 and Table A.4). We obtained
262 unpublished full multivariable model data relating to the trial of Björnsson Hallgren et
263 al.[40]
264
265 *Included studies*
266 Study characteristics
267 Key characteristics of the five studies are presented in Table 1. The studies were
268 published between 2005 and 2014. All appeared to have been conducted in
269 outpatient settings. All were cohort studies, but in two the cohort was derived from
270 pooled data from an RCT.[39,40] None of the studies was prospectively registered;
271 however, the intention for a prognostic investigation was mentioned in the published
272 protocol[44] for the study by Kromer et al. 2014. Four studies concerned model
273 development and the fifth[42] was reported as a validation study.
274
275 Four studies[39-41,43] investigated mixed populations with impingement-related shoulder
276 pain. One of these[41] excluded FTT. One study[42] investigated a rotator cuff tear
277 population without differentiating between PTT and FTT. Initial sample sizes ranged
278 from 33[41] to 102;[40,43] with the number of outcome events (number of patients in
279 whom the prognosticated event occurred) ranging from 23[41] to 89.[43] Although varying
280 in duration, content and dosage, physiotherapy was provided to all study participants;
281 steroid injections were provided to all participants of one study[40] and were optional in
282 another.[43]
283

The number of initially considered prognostic factors was unclear in three studies,[39-41] but, based on the presented data, appeared to range from eight[43] to presumably over 60.[41] Prognostic factors mainly involved demographics and clinical characteristics such as symptoms or diagnostic imaging findings. One study[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.[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.[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[41] to 12 months.[39,40,42,43]

The methods for selecting prognostic factors for multivariable analysis, where specified, varied (Table A.5); two studies[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.[39,41,43] The nominal validation study by Merolla[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[39,42,43]; unclear or lack of blinding of outcome determination to prognostic factor information;[41,42] 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);[40,42] unclear handling of missing data;[39-43] use of univariable analyses to select prognostic factors;[39,41] unclear[40] or unspecified[42] modelling methods; and failure to consider overfitting of data, complexities in the data, evaluation of performance measures or non-linear relationships.[39-43]

Overall concerns about applicability mainly related to domain 2 (predictors) and were rated as low for two studies,[39,40] unclear for one,[43] and high for two.[41,42] The high concerns related to the definition and assessment of prognostic factors in two studies.[41,42] 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.[41] provided a prognostic index (Table A.5).

Conflicts of interest

Conflicts of interest were explicitly addressed only in two studies,[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[39-41,43] or validate[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[39-41,43] investigated populations with impingement-related shoulder pain, implicitly including rotator cuff tears of varying completeness, except Hung et al,[41] who excluded FTT. Merolla et al.[42] 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[41] or rotator cuff tears.[39,40,43] Also, in two studies the patient populations were selected, by dint of their agreement to participate in an RCT,[39,40] which may have reduced external validity.

The physiotherapy intervention was insufficiently described to allow a judgement in Taheriazam et al.[43] However, in the intervention group of Björnsson-Hallgren et al.[40] and in the other three studies,[39,41,42] the physiotherapy intervention was generally consistent with standard practice.[45,46]

Less uniform was the selection of predictors, which was generally unjustified and diverse. In one case,[41] 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.[41] 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[42], shoulder disability[39,42,43] and Global Perceived Change.[41] Björnsson-Hallgren[40]

388 reported the decision to undergo surgery. The remaining outcomes of interest for this
389 review, including adverse events, HRQoL, sick leave and structural progression of
390 tears, were either not reported or, in one case,[42] reported too unclearly for extraction.
391
392 None of the four development studies[39-41,43] reported any form of internal model
393 validation; and none of these four was followed by an external validation, even
394 though five and 10 years had elapsed since Hung et al.[41] and Taheriazam et al.[43]
395 respectively. Lack of appropriate validation of prognostic models is a widely observed
396 issue.[47] There is good empirical evidence that models perform substantially less well
397 in external, i.e. independent, samples, and that performance in external samples is
398 more representative of clinical performance,[28,48] so this presents a major obstacle to
399 usability. The fifth study (Merolla et al.),[42] though reportedly a validation, was
400 seriously flawed in both concept and execution. Ultimately none of the studies has
401 been assessed for clinical impact and, consequently, none of the models presented
402 in the included studies is usable in clinical practice.
403
404 *Quality of the evidence*
405 We evaluated risk of bias in five domains: participant selection, predictors, outcome,
406 sample size and flow, and analysis. Our judgment of risk of bias was affected by a
407 number of methodological issues (see results). Most of the identified deficiencies
408 have been addressed extensively in the literature; several, including in particular
409 those relating to the number of prognostic factors in relation to the number of
410 outcome events and use of univariable analyses to select prognostic factors have
411 been shown to result in invalid and unreliable models.[49] Similarly, the use of
412 statistical methods such as stepwise regression to select factors within the
413 multivariable analysis has been criticized.[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.,[42] 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.[51] Mathematical coupling, which inherently occurs "when one variable directly or indirectly contains the whole or part of another"[52] 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.[53] Moreover, in the present context, PROMs are among very few prognostic factors that have a basis in evidence[37,38] In our review, this conflict was encountered in two studies, Kromer et al.[39] and Taheriazam et al.,[43] 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[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,[e.g.34,55,56] but had concerns about the currency of all but one,[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.[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

466 referred to the latest recommendations of the PROGRESS partnership[25] and, after
467 piloting earlier versions, evaluated risk of bias and applicability using a near-definitive
468 but unpublished version of PROBAST (R. Wolff, personal communication). The use
469 of PROBAST was especially appealing to us because it is the first tool to specifically
470 address risk of bias and applicability in prognostic model studies.
471
472 *Agreements and disagreements with other studies or reviews*
473 To our knowledge, this is the first systematic review to synthesize the evidence on
474 primary prognostic model research in adults with rotator cuff disorders who are
475 undergoing conservative treatment with physiotherapy. We identified two other
476 prognostic systematic reviews addressing shoulder pain,[37,38] but both aimed to
477 synthesize evidence on individual prognostic factors rather than on prognostic
478 models, and have minimal overlap with our own review, which has a single study[41] in
479 common with Chester et al.[37] and none with Kuijpers et al.[38] Of the two reviews,
480 Chester et al.,[37] like us, limited inclusion to studies investigating response to
481 conservative treatment with physiotherapy, while Kuijpers et al.[38] studied overall
482 prognosis. Both reviews addressed shoulder pain in general and did not provide any
483 subgroup analyses to allow for inferences about rotator cuff disorders. Thus, while
484 evidence was found supporting a limited number of emerging factors including
485 symptom duration, baseline function or disability,[37,38] pain and age,[38] the
486 transferability of these findings to the population of interest in our review is unclear.
487
488 **Conclusions**
489 *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,[35] 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.[57] 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.[58]

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

516 on their use. We also acknowledge the support from Iain Baird and Julie Hogg,
517 Library and Information Services, Teesside University, Middlesbrough, UK, on the
518 development of the search strategy; and Hanna Björnsson Hallgren for providing
519 unpublished analysis data.

520

521

522 **Disclosure of potential conflicts of interest:**

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

References

[1] Kooijman M, Swinkels I, van Dijk C, et al. Patients with shoulder syndromes in general and physiotherapy practice: an observational study. *BMC Musculoskelet Disord* 2013;14:128.

[2] Paloneva J, Koskela S, Kautiainen H, et al. Consumption of medical resources and outcome of shoulder disorders in primary health care consulters. *BMC Musculoskelet Disord* 2013;14:348.

[3] Virta L, Joranger P, Brox JI, et al. Costs of shoulder pain and resource use in primary health care: a cost-of-illness study in Sweden. *BMC Musculoskelet Disord* 2012;13:17.

[4] Östör AJK, Richards CA, Prevost AT, et al. Diagnosis and relation to general health of shoulder disorders presenting to primary care. *Rheumatology* 2005;44:800–5.

[5] Van der Windt DA, Koes BW, de Jong BA, et al. Shoulder disorders in general practice: incidence, patient characteristics, and management. *Ann Rheum Dis* 1995;54:959–64. DOI:10.1136/ard.54.12.959.

[6] Cook JL, Purdam CR. Is tendon pathology a continuum? A pathology model to explain the clinical presentation of load-induced tendinopathy. *Br J Sports Med* 2009;43:409–16.

[7] Reilly P, Macleod I, Macfarlane R, et al. Dead men and radiologists don't lie: a review of cadaveric and radiological studies of rotator cuff tear prevalence. *Ann R Coll Surg Engl* 2006;88:116–21.

[8] Beaudreuil J, Bardin T, Orcel P, et al. Natural history or outcome with conservative treatment of degenerative rotator cuff tears. *Jt Bone Spine* 2007;74:527–9.

[9] Hanchard NCA, Lenza M, Handoll HHG, et al. Physical tests for shoulder impingements and local lesions of bursa, tendon or labrum that may accompany impingement. *Cochrane Database Syst Rev* 2013;4:CD007427.

[10] Minns Lowe CJ, Moser J, Barker K. Living with a symptomatic rotator cuff tear “bad days, bad nights”: a qualitative study. *BMC Musculoskelet Disord* 2014;15:228.

[11] Piitulainen K, Ylinen J, Kautiainen H, et al. The relationship between functional disability and health-related quality of life in patients with a rotator cuff tear. *Disabil Rehabil* 2012;34:2071–5.

[12] Lenza M, Buchbinder R, Takwoingi Y, et al. Magnetic resonance imaging, magnetic resonance arthrography and ultrasonography for assessing rotator cuff tears in people with shoulder pain for whom surgery is being considered. *Cochrane Database Syst Rev* 2013;9:CD009020.

[13] American Academy of Orthopaedic Surgeons (AAOS). American Academy of Orthopaedic Surgeons clinical practice guideline: optimizing the management of rotator cuff problems, www.aaos.org/research/guidelines/RCP_guideline.pdf (2010, accessed 21 June 2015).

[14] Beaudreuil J, Dhénain M, Coudane H, et al. Clinical practice guidelines for the surgical management of rotator cuff tears in adults. *Orthop Traumatol Surg Res* 2010;96:175–9.

[15] Diercks R, Bron C, Dorrestijn O, et al. Guideline for diagnosis and treatment of subacromial pain syndrome. *Acta Orthop* 2014;85:314–22.

[16] Ketola S, Lehtinen J, Rousi T, et al. No evidence of long-term benefits of arthroscopic acromioplasty in the treatment of shoulder impingement syndrome: Five-year results of a randomised controlled trial. *Bone Joint Res* 2013;2:132–9.

[17] Kukkonen J, Joukainen A, et al. Treatment of non-traumatic rotator cuff tears: A randomised controlled trial with one-year clinical results. *Bone Jt J* 2014;96 B:75–81.

[18] Moosmayer S, Lund G, Seljom US, et al. Tendon Repair Compared with Physiotherapy in the Treatment of Rotator Cuff Tears: A Randomized

Controlled Study in 103 Cases with a Five-Year Follow-up. *J Bone Jt Surg* 2014;96:1504–14.

[19] Saltychev M, Äärimaa V, Virolainen P, et al. Conservative treatment or surgery for shoulder impingement: systematic review and meta-analysis. *Disabil Rehabil* 2014:1–8.

[20] Colvin AC, Egorova N, Harrison AK, et al. National trends in rotator cuff repair. *J Bone Joint Surg Am* 2012;94:227–33.

[21] Svendsen S, Frost P, Jensen L. Time trends in surgery for non-traumatic shoulder disorders and postoperative risk of permanent work disability: a nationwide cohort study. *Scand J Rheumatol* 2012;41:59–65.

[22] Butler M, Forte M, Braman J, et al. Nonoperative and Operative Treatments for Rotator Cuff Tears: Future Research Needs. Future Research Needs Paper No. 39. AHRQ Publication No. 13-EHC050-EF. Rockville, MD: Agency for Healthcare Research and Quality (US), www.effectivehealthcare.ahrq.gov/reports.final.cfm (2013, accessed 21 June 2015).

[23] Croft P, Altman DG, Deeks JJ, et al. The science of clinical practice: disease diagnosis or patient prognosis? Evidence about “what is likely to happen” should shape clinical practice. *BMC Med* 2015;13:20.

[24] Moons KGM, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: what, why, and how? *BMJ* 2009;338:b375.

[25] PROGRESS, MRC PROGnosis RESearch Strategy Partnership, www.progress-partnership.org/ (2015, accessed 21 June 2015).

[26] Cochrane Prognosis Methods Group, prognosismethods.cochrane.org/ (2015, accessed 21 June 2015).

[27] Hemingway H, Croft P, Perel P, et al. Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. *BMJ* 2013;346:e5595.

[28] Steyerberg EW, Moons KGM, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. *PLoS Med* 2013;10:e1001381.

[29] Galea MH, Blamey RW, Elston CE, et al. The Nottingham prognostic index in primary breast cancer. *Breast Cancer Res Treat* 1992;22:207–19.

[30] Harrell FE. *Regression modeling strategies*. 1st ed. Springer New York; 2001.

[31] PROSPERO, International Prospective Register of Systematic Reviews, www.crd.york.ac.uk/PROSPERO/ (2015, accessed 21 June 2015).

[32] Braun C, Bularczyk M, Heintsch J, et al. Manual therapy and exercises for shoulder impingement revisited. *Phys Ther Rev* 2013;18:263–84.

[33] Braun C, Hanchard NCA. Manual therapy and exercise for impingementrelated shoulder pain. *Phys Ther Rev* 2010;15:62–83.

[34] Geersing GJ, Bouwmeester W, Zuithoff P, et al. Search filters for finding prognostic and diagnostic prediction studies in medline to enhance systematic reviews. *PLoS One* 2012;7:e32844.

[35] PROBAST, Prediction Risk of Bias Tool 2015. York, UK: Kleijnen Systematic Reviews Ltd., www.systematic-reviews.com/probast (2015, accessed 21 June 2015).

[36] QUADAS-2 2015. University of Bristol, UK; www.bris.ac.uk/quadas/ (2015, accessed 21 June 2015).

[37] Chester R, Shepstone L, Daniell H, et al. Predicting response to physiotherapy treatment for musculoskeletal shoulder pain: a systematic review. *BMC Musculoskelet Disord* 2013;14:203.

[38] Kuijpers T, Van Der Windt DAWM, Van Der Heijden GJMG, et al. Systematic review of prognostic cohort studies on shoulder disorders. *Pain* 2004;109:420–31.

[39] Kromer TO, Sieben JM, de Bie RA, et al. Influence of Fear-Avoidance Beliefs on Disability in Patients With Subacromial Shoulder Pain in Primary Care: A Secondary Analysis. *Phys Ther* 2014;94:1175–84.

[40] Björnsson Hallgren HC, Holmgren T, Oberg B, et al. A specific exercise strategy reduced the need for surgery in subacromial pain patients. *Br J Sports Med* 2014.

[41] Hung CJ, Jan MH, Lin YF, et al. Scapular kinematics and impairment features for classifying patients with subacromial impingement syndrome. *Man Ther* 2010;15:547–51.

[42] Merolla G, Paladini P, Saporito M, et al. Conservative management of rotator cuff tears: literature review and proposal for a prognostic. Prediction Score. *Muscles Ligaments Tendons* J 2011;1:12–9.

[43] Taheriazam A, Sadatsafavi M, Moayyeri A. Outcome predictors in nonoperative management of newly diagnosed subacromial impingement syndrome: a longitudinal study. *Med Gen Med* 2005; 7:63.

[44] Kromer TO, de Bie RA, Bastiaenen CHG. Effectiveness of individualized physiotherapy on pain and functioning compared to a standard exercise protocol in patients presenting with clinical signs of subacromial impingement syndrome. A randomized controlled trial. *BMC Musculoskelet Disord* 2010;11:114.

[45] Struyf F, De Hertogh W, Gulinck J, et al. Evidence-based treatment methods for the management of shoulder impingement syndrome among dutch-speaking physiotherapists: An online, web-based survey. *J Manipulative Physiol Ther* 2012;35:720–6.

[46] Varela E, Valero R, Küçükdeveci AA, et al. Shoulder pain management. The role of physical and rehabilitation medicine physicians. The European perspective based on the best evidence. A paper by the UEMS-PRM section professional practice committee. *Eur J Phys Rehabil* Med 2013;49:743–51.

[47] Steyerberg EW, Harrell FE. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol*. Epub ahead of print 18 April 2015. DOI: 10.1016/j.jclinepi.2015.04.005.

[48] Siontis GC, Tzoulaki I, Castaldi PJ, et al. External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination. *J Clin Epidemiol* 2015;68:25–34.

[49] Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361–87.

[50] Flom PL. Stopping stepwise : Why stepwise and similar selection methods are bad, and what you should use. NESUG, www.nesug.org/proceedings/nesug07/sa/sa07.pdf (2007, accessed 22 March 2015).

[51] Reitsma JB, Rutjes AWS, Whiting P, et al. Chapter 9: Assessing methodological quality. In: Deeks J.J., Bossuyt PM, Gatsonis, C. (editors). *Cochrane Handbook of systematic reviews of diagnostic test accuracy* version 1.0.0., srdta.cochrane.org/ (2009, accessed 21 June 2015).

[52] Tu YK, Gilthorpe MS. Revisiting the relation between change and initial value: A review and evaluation. *Stat Med* 2007;26:443–57.

[53] Vodicka E, Kim K, Devine EB, et al. Inclusion of Patient-Reported Outcome Measures in Registered Clinical Trials: Evidence from ClinicalTrials.gov (2007-2013). *Contemp Clin Trials*. Epub ahead of print 18 April 2015. DOI: 10.1016/j.cct.2015.04.004.

[54] Chatterley T, Dennett L. Utilisation of search filters in systematic reviews of prognosis questions. *Health Info Libr J* 2012;29:309–22.

[55] Altman DG. Systematic reviews of evaluations of prognostic variables. BMJ 2001;323:224–8.

[56] Walker-Dilks C, Wilczynski NL, Haynes RB. Cumulative Index to Nursing and Allied Health Literature search strategies for identifying methodologically sound causation and prognosis studies. *Appl Nurs Res* 2008;21:98–103.

[57] Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015;350:g7594–g7594.

[58] Moons KG, de Groot JA, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: The CHARMS Checklist. *PLoS Med* 2014;11:e1001744.

[59] Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions explanation and elaboration. *Br Med J* 2009:b2700.

List of tables and figures

List of tables

Tables as supplementary materials:

List of figures

Table 1. Characteristics of included studies (alphabetical order)

	Björnsson Hallgren 2014
DESIGN	Cohort study derived from 2 group RCT; model development
SETTING	Sweden; presumably outpatient
STARTPOINT	Not precisely defined; recruitment was from the waiting list for arthroscopic subacromial decompression (duration of symptoms ≥ 6 months)
PARTICIPANTS	N = 102 (data on 95)*; "subacromial pain"; mixed population: non-tear (69%), partial tear (22%), full tear (9%)
INTERVENTION	Exercise-based physiotherapy for both groups (specific versus control exercises) following initial steroid injection (duration 12 weeks)
PROGNOSTIC FACTORS CONSIDERED	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)
OUTCOME	Choice of surgery (yes/no, based on record of treatment) OUTCOME EVENTS n = 41
ENDPOINT	After 1 year (after inclusion or after surgery)
SELECTION OF FACTORS	FOR MULTIVARIABLE MODELLING unclear (lack of information); there is no suggestion of predictor selection based on univariable analysis. WITHIN MULTIVARIABLE MODELLING unclear (lack of information). No rationale was provided for the combinations of prognostic factors, and no 'final' model was specified, but apparently, no stepwise regression was used.
STATISTICAL ANALYSIS	Logistic regression
NOTES	*Unpublished analysis data specifies up to 97 observations

	Hung 2010
DESIGN	Cohort (single-group); developmental; model development
SETTING	Taiwan; presumably outpatient
STARTPOINT	Recruitment by an orthopaedics clinic or by "general announcements in the local internet media"; no further information was provided
PARTICIPANTS	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)
INTERVENTION	Standardized 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, with dichotomisation into "improved" or "not improved". OUTCOME EVENTS n = 23
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	Logistic regression

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

	Kromer 2014
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

	outcomes in the Methods. OUTCOME EVENTS 33 for continuous outcomes (conservatively treated participants)
ENDPOINT	Unclear. Outcomes were measured at 6, 9 & 12 months, but the prognosis aspect may have been assessed at 12 months only.
STATISTICAL ANALYSIS	Unclear. "Student's t-test was used to highlight significant differences between pre- and post-rehabilitation program scores."
NOTES	[§]Unclear & incomplete reporting seriously hindered data extraction.

	Taheriazam 2005
DESIGN	Cohort (single-group); model development
SETTING	Iran; outpatient
STARTPOINT	New diagnosis of impingement syndrome
PARTICIPANTS	N = 102[‖] (data for 89); "subacromial impingement syndrome" (NI on whether or not rotator cuff tears were included)
INTERVENTION	Treatment was based on a standardised protocol including oral NSAIDs, up to two local steroid injections and a supervised physical therapy program (overall duration presumably 12 months)
PROGNOSTIC FACTORS CONSIDERED	N = 8 Acromial morphology (type I, II or III)[F], 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")[F], age, Constant score, dominant shoulder involvement (yes or no), sex, symptom duration.
OUTCOMES	Constant score OUTCOME EVENTS n = 89
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 ($p<0.05$). Among the three remaining predictors, three further multivariable models were then calculated.
STATISTICAL ANALYSIS	Linear regression
NOTES	[‖]As reported by the authors, but there is a discrepancy. Of 128 eligible patients, 93 consented & 13 were excluded from the analysis, giving a sample of 80. [F]Categorized but erroneously analysed as continuous data

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, ROM = Range of Motion, SD = Standard Deviation, SLAP = Superior Labral Anterior to Posterior, SPADI-F = Shoulder Pain & Disability Index Function subscale, VAS = Visual Analogue Scale, VNRS = Visual Numeric Rating Scale.

Table 2: PROBAST (risk of bias and applicability) ratings

Study	RISK OF BIAS				APPLICABILITY CONCERNS			OVERALL JUDGEMENTS			
	1. PARTICIPANT SELECTION	2. PREDICTORS	3. OUTCOME	4.SAMPLE SIZE & FLOW	5. ANALYSIS	1. PARTICIPANT SELECTION	2. PREDICTORS	3. OUTCOME	RISK OF BIAS	APPLICABILITY	USABILITY OF THE MODEL
Björnsson Hallgren 2014	☺	☺	☺	☹	?	☺	☺	☺	☹	☺	☹
Hung 2010	☺	?	☹	?	☹	?	☹	?	☹	☹	☹
Kromer 2014	☺	☺	☹	?	☹	☺	☺	☺	☹	☺	☹
Merolla 2011	?	?	☹	☹	?	?	☹	?	☹	☹	☹
Taheriazam 2005	☺	☺	☹	?	?	?	?	☺	☹	?	☹

☺ Low risk/concerns ☹ High risk/concerns ? Unclear risk/concerns

Table A.1: Deviations from protocol

Aspect	Difference with justification
Outcome	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.
Presentation of prognostic model	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.
Author contact	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.

Table A.2: Medline search strategy (EBSCO format)

S1	((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”
S2	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
S3*	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*))
S4*	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
S5†	prognos*.ti,ab
S6	S1 AND S2 AND (S3 OR S4 OR S5)

* 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 A.3: Excluded studies (alphabetically ordered by first author)

Nr	Study	Main reasons for exclusion – criterion categories* (in brackets) and explanations
1	**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.	(Po) Traumatic population (all participants had experienced a "posttraumatic rotator cuff tear in an industrial accident")
2	**Bartolozzi** A, Andreychik D, Ahmad S. Determinants of outcome in the treatment of rotator cuff disease. Clin Orthop Relat Res 1994:90–7.	(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
3	**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.	(Po) Only 24% were atraumatic (S/A) No multivariable prognostic modelling; retrospective
4	**Bonde** JP, Mikkelsen S, Andersen JH, Fallentin N, Baelum J, Svendsen SW, et al. Prognosis of shoulder tendonitis in repetitive work: a follow up study in a cohort of Danish industrial and service workers. Occup Environ Med 2003;60:E8.	(I) No mention of physiotherapy (not following a defined course of conservative treatment with physiotherapy)
5	Boorman RS, More KD, Hollinshead RM, Wiley JP, Brett K, Mohtadi NG, et al. The rotator cuff quality-of-life index predicts the outcome of nonoperative treatment of patients with a chronic rotator cuff tear. J Bone Joint Surg Am 2014;96:1883–8. doi:10.2106/JBJS.M.01457.	(Po) In 49%, the onset was traumatic.
6	**Chard** MD, Sattelle LM, Hazleman BL. The long-term outcome of rotator cuff tendinitis--a review study. Br J Rheumatol 1988;27:385–9.	(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
7	**Contreras** F, Brown HC, Marx RG. Predictors of success of corticosteroid injection for the management of rotator cuff disease. HSS J Musculoskelet J Hosp Spec Surg 2013;9:2–5. doi:10.1007/s11420-012-9316-6.	(S/A) No multivariable prognostic modelling
8	**Cummins** CA, Sasso LM, Nicholson D. Impingement syndrome: Temporal outcomes of nonoperative treatment. J Shoulder Elb Surg 2009;18:172–7. doi:10.1016/j.jse.2008.09.005.	(S/A) No multivariable prognostic modelling Despite the allusion to logistic regression analysis (p. 173), no multivariable analysis is reported in the results. Author contact failed to resolve this issue.
9	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.	(S/A) Cross-sectional study; no multivariable prognostic modelling

	doi:10.1177/2325967115584596.	
10	**Deutscher** D, Horn SD, Dickstein R, Hart DL, Smout RJ, Gutvirtz M, et al. Associations Between Treatment Processes, Patient Characteristics, and Outcomes in Outpatient Physical Therapy Practice. Arch Phys Med Rehabil 2009;90:1349–63. doi:10.1016/j.apmr.2009.02.005.	(Po) Population not condition-specific (shoulder pain as one out of four musculoskeletal impairment groups categories)
11	**Ekeberg** OM, Bautz-Holter E, Juel NG, Engebretsen K, Kvalheim S, Brox JI. Clinical, socio-demographic and radiological predictors of short-term outcome in rotator cuff disease. BMC Musculoskelet Disord 2010;11:239. doi:10.1186/1471-2474-11-239.	(I) No defined physiotherapy treatment: was allowed if started, but was not followed. Only some patients had physiotherapy (see primary RCT report)
12	**Engebretsen** K, Grotle M, Bautz-Holter E, Ekeberg OM, Brox JI. Predictors of shoulder pain and disability index (SPADI) and work status after 1 year in patients with subacromial shoulder pain. BMC Musculoskelet Disord 2010;11:218. doi:10.1186/1471-2474-11-218.	(S/A) Secondary, retrospective analysis (no mention of planned prognostic analysis in trial registry entry (NCT00653081))
13	Ertan S, Ayhan E, Güven MF, Kesmezacar H, Akgün K, Babacan M. Medium-term natural history of subacromial impingement syndrome. J Shoulder Elbow Surg 2015:1512–8. doi:10.1016/j.jse.2015.06.007.	(S/A) Retrospective study; no multivariable prognostic modelling (I) Not investigating a course of conservative treatment with physiotherapy (reference is made to initial treatment consisting of medication, cold compression and modification of activities)
14	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:10.1177/2325967114S00059.	(TP) Extended abstract; no published full study report available (may be linked with Kweon et al. 2015)
15	**Gialanella** B, Bertolinelli M. Corticosteroids injection in rotator cuff tears in elderly patient: Pain outcome prediction. Geriatr Gerontol Int 2013;13:993–1001. doi:10.1111/ggi.12046.	(Po) All patients had some degree of shoulder osteoarthritis
16	**Goldberg** BA, Nowinski RJ, Matsen FA. Outcome of nonoperative management of full-thickness rotator cuff tears. Clin Orthop Relat Res 2001;(382):99–107.	(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
17	**Hardy** DC, Vogler JB, White RH. The shoulder impingement syndrome: prevalence of radiographic findings and correlation with response to therapy. Am J Roentg 1986;147;3:557-61 doi:10.1016/0002-9343(86)90696-0.	(I) No involvement of physiotherapy (S/A) No multivariable prognostic modelling

18	**Hawkins** RH, Dunlop R. Nonoperative treatment of rotator cuff tears. Clin Orthop Relat Res 1995:178–88.	(Po) 64% of cases were traumatic (O) The outcome variable is patient satisfaction, which is not an outcome of interest in this review
19	**Itoi** E, Tabata S. Conservative treatment of rotator cuff tears. Clin Orthop Relat Res 1992:165–73.	(S/A) No multivariable prognostic modelling; possibly retrospective
20	**Kaergaard** A, Andersen JH. Musculoskeletal disorders of the neck and shoulders in female sewing machine operators: prevalence, incidence, and prognosis. Occup Environ Med 2000;57:528–34. doi:10.1136/oem.57.8.528.	(I) Not following a course of conservative treatment with physiotherapy
21	**Kennedy** CA, Haines T, Beaton DE. Eight predictive factors associated with response patterns during physiotherapy for soft tissue shoulder disorders were identified. J Clin Epidemiol 2006;59:485–96. doi:10.1016/j.jclinepi.2005.09.003.	(Po) Generic shoulder soft-tissue disorder population with no distinct impingement spectrum subgroup
22	**Kennedy** CA, Manno M, Hogg-Johnson S, Haines T, Hurley L, McKenzie D, et al. Prognosis in soft tissue disorders of the shoulder: predicting both change in disability and level of disability after treatment. Phys Ther 2006;86:1013–32.	Duplicate publication of Kennedy 2006a (see above)
23	Ketola S, Lehtinen J, Rousi T, Nissinen M, Huhtala H, Arnala I. Which patients do not recover from shoulder impingement syndrome, either with operative treatment or with nonoperative treatment? Acta Orthop 2015;86:1–6. doi:10.3109/17453674.2015.1033309.	(S/A) Secondary, retrospective analysis (no mention of planned prognostic analysis in trial registry entry (NCT00349648))
24	**Kijima** H, Minagawa H, Nishi T, Kikuchi K, Shimada Y. Long-term follow-up of cases of rotator cuff tear treated conservatively. J Shoulder Elb Surg 2012;21:491–4. doi:10.1016/j.jse.2011.10.012.	(S/A) No multivariable prognostic modelling
25	**Kulenkampff** H-A, Reichelt A. Clinical course of ruptures of the rotator cuff after conservative therapy. Orthopadische Prax 1990;26:493–6.	(L) Full text in German (S/A) Not a prognostic model study
26	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.	(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.
27	**Laslett** M, Steele M, Hing W, McNair P, Cadogan A. Shoulder pain in primary care - Part 2: Predictors of clinical outcome to 12 months. J Rehabil Med 2014. [Epub ahead of print]. doi:10.2340/16501977-1885.	(Po) Mixed shoulder pain population, no subgroup analyses for rotator cuff disorders (I) No follow-up of a course of physiotherapy (physiotherapy treatment was not documented)

28	**Maman** E, Harris C, White L, Tomlinson G, Shashank M, Boynton E. 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.	(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
29	**McCreesh** K. Evidence based prognosis setting in the case of a conservatively managed rotator cuff tear. Physiother Irel 2007;28:31–5.	(S/A) Case study
30	**Mintken** PE, Cleland JA, Carpenter KJ, Bieniek ML, Keirns M, Whitman JM. Some factors predict successful short-term outcomes in individuals with shoulder pain receiving cervicothoracic manipulation: a single-arm trial. Phys Ther 2010;90:26–42. doi:10.2522/ptj.20090095.	(Po) Generic shoulder pain population; 30% due to trauma
31	**Morag** Y, Jamadar DA, Miller B, Brandon C, Gandikota G, Jacobson JA. 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.	(I) 80% of the sample received physiotherapy, but these are not separately reported (S/A) No multivariable prognostic modelling
32	Nakamura Y, Yokoya S, Mochizuki Y, Harada Y, Kikugawa K, Ochi M. Monitoring of progression of nonsurgically treated rotator cuff tears by magnetic resonance imaging. J Orthop Sci 2015;20:314–20. doi:10.1007/s00776-014-0680-6.	(S/A) No multivariable prognostic modelling
33	Notarnicola A, Maccagnano G, Tafuri S, Fiore A, Margiotta C, Pesce V, et al. Prognostic factors of extracorporeal shock wave therapy for tendinopathies. Musculoskelet Surg 2015. doi:10.1007/s12306-015-0375-y.	(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
34	**Ottaviani** M, Mele G. Epidemiological, clinical and diagnostic study of rotator cuff rupture. Riabilitazione 1998;31:17–24.	(L) Full text in Italian (S/A) Presumably anyway not a multivariable prognostic modelling study
35	**Rahme** H, Solem-Bertoft E, Westerberg CE, Lundberg E, Sörensen S, Hilding S. 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.	(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)
36	**Rowe** CR. Ruptures of the rotator cuff: selection of cases for conservative treatment..	(S/A) Not a prognostic modelling study

	Surg Clin North Am 1963;43:1531–4.	
37	**Ryall** C, Coggon D, Peveler R, Poole J, Palmer KT. A prospective cohort study of arm pain in primary care and physiotherapy--prognostic determinants. Rheumatology (Oxford) 2007;46:508–15. doi:10.1093/rheumatology/kel320.	(P) Non-specific population ("arm pain"), no sub-classification of shoulder pain
38	**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.	(Po) 53% were post-traumatic (I) No mention of physiotherapy; not following a defined course of conservative treatment with physiotherapy ("natural progression") (S/A) There appears to be no prognostic modelling
39	**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.	(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
40	**Silverstein** BA, Viikari-Juntura E, Fan ZJ, Bonauto DK, Bao S, Smith C. Natural course of nontraumatic rotator cuff tendinitis and shoulder symptoms in a working population. Scand J Work Environ Heal 2006;32:99–108. doi:10.5271/sjweh.985.	(I) The proportion receiving physiotherapy is not specified; not following a defined course of conservative treatment with physiotherapy ("natural course") (S/A) No multivariable prognostic modelling
41	**Sindhu** BS, Lehman LA, Tarima S, Bishop MD, Hart DL, Klein MR, et al. Influence of Fear-Avoidance Beliefs on Functional Status Outcomes for People With Musculoskeletal Conditions of the Shoulder. Phys Ther 2012;92:992–1005. doi:10.2522/ptj.20110309.	(P) ICD-9 classifications (disease categories) too imprecise for localisation to the rotator cuff disorder spectrum as defined for this review (S) Retrospective study
42	**Smith** KL, Harryman DT, Antoniou J, Campbell B, Sidles JA, Matsen FA. A prospective, multipractice study of shoulder function and health status in patients with documented rotator cuff tears. J Shoulder Elb Surg 2000;9:395–402. doi:10.1067/mse.2000.108962.	(S) Not a prognostic modelling study: effectively a time zero cross sectional analysis correlating various clinical characteristics with Simple Shoulder Test (SST) functions
43	**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.	(I) Treatment unspecified (not all patients received physiotherapy); i.e. not following a defined course of conservative treatment with physiotherapy (O) Outcome of interest ("referral" to a secondary care specialist) not of interest for this review
44	**Tanaka** M, Itoi E, Sato K, Hamada J, Hitachi S, Tojo Y, et al. Factors related to successful outcome of conservative treatment for rotator cuff tears. Ups J Med Sci 2010;115:193–200. doi:10.3109/03009734.2010.493246.	(S/A) No multivariable prognostic modelling
45	**Van Der Windt** DAWM, Koes BW, Boeke AJP, Devillé W, De Jong B a, Bouter LM. Shoulder	(I) Not following a defined course of conservative treatment with physiotherapy:

	disorders in general practice: Prognostic indicators of outcome. Br J Gen Pract 1996;46:519–23.	not all patients (in the rotator cuff tendinitis group) had physiotherapy
46	**Viikari-Juntura** E, Takala EP, Riihimäki H, Martikainen R, Jäppinen P. Predictive validity of symptoms and signs in the neck and shoulders. J Clin Epidemiol 2000;53:800–8. doi:10.1016/S0895-4356(00)00197-9.	(Po) Non-specific shoulder pain population (I) Physiotherapy not for all participants
47	**Virta** L, Mortensen M, Eriksson R, Möller M. How many patients with subacromial impingement syndrome recover with physiotherapy? A follow-up study of a supervised exercise programme. Adv Physiother 2009;11:166–73. doi:10.1080/14038190802460481.	(S/A) No multivariable prognostic modelling
48	**Wang** JC, Horner G, Brown ED, Shapiro MS. The relationship between acromial morphology and conservative treatment of patients with impingement syndrome. Orthopedics 2000;23:557–9.	(S) No multivariable prognostic modelling
49	**Yamanaka** K, Matsumoto T. The joint side tear of the rotator cuff. A followup study by arthrography. Clin Orthop Relat Res 1994;304:68-73. doi:10.1097/00003086-199407000-00012.	(Po) 28% post-traumatic (I) Conservative treatment is undefined (S/A) No multivariable prognostic modelling

*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

Study	Source
ICTRP (study ID and title), ordered by ID	
ACTRN12615000351516 Pain modulation characteristics in people with shoulder impingement and predictors of successful outcomes following physiotherapy treatment	ICTRP. Available at: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=ACTRN12615000351516 [last accessed 22 Oct 2015]
DRKS00004462 Predicting the outcome of conservative treatment with physiotherapy for shoulder pain in the presence of atraumatic partial-thickness tears of the rotator cuff	ICTRP. Available at: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00004462 [last accessed 22 Oct 2015]
NCT00632996 Exercise and Manual Therapy for Shoulder Subacromial Impingement Syndrome	ICTRP. Available at: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00632996 [last accessed 22 Oct 2015]
NCT00762580 Features to Predict Success With Nonoperative Treatment of Patients With Rotator Cuff Tears (MOON)	ICTRP, Available at: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00762580 [last accessed 22 Oct 2015]
NCT01498198 Workers Compensation Board: Rotator Cuff Tear Management	ICTRP. Available at: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01498198 [last accessed 22 Oct 2015]
NCT02287090 Comparative Effectiveness of Operative Versus Non-Operative Treatments for Rotator Cuff Tears (ROW)	ICTRP. Available at: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02287090 [last accessed 22 Oct 2015]
NCT02510352 Cohort of Patients With a Symptomatic Rotator Cuff Tear Treated Without Surgical Repair	ICTRP. Available at: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02510352 [last accessed 22 Oct 2015]
Published protocol (first author (year))	
Lambers Heerspink (2011) Clinical and radiological outcome of conservative vs. surgical treatment of atraumatic degenerative rotator cuff rupture: design of a randomized controlled trial.	Lambers Heerspink FO, Hoogeslag Ag R, Diercks L R, van Eerden PJ, van den Akker-Scheek I, van Raay JJ. Clinical and radiological outcome of conservative vs. surgical treatment of atraumatic degenerative rotator cuff rupture: design of a randomized controlled trial. BMC Musculoskelet Disord 2011;12:25.doi: 10.1186/1471-2474-12-25

Table A.5 Characteristics and results of included studies – detailed version

	Björnsson Hallgren 2014
OBJECTIVE	To examine ... „whether the baseline Constant-Murley (CM) score, rotator cuff status and radiological findings influenced the decision about surgery.“
DESIGN	Cohort study derived from 2 group RCT; consecutive recruitment **PHASE OF RESEARCH** development
SETTING	Sweden; orthopaedic [presumably outpatient] department in a university hospital. **STUDY DATES** recruitment took place from January 2008 to February 2010.
STARTPOINT	Not precisely defined; recruitment was from the waiting list for arthroscopic subacromial decompression (duration of symptoms ≥ 6
PARTICIPANTS	**N** 102 (data on 95)* **TYPE OF DISORDER** “subacromial pain”**;** mixed population: non-tear (69%), partial tear (22%), full tear (9%) **MEAN AGE** 52 years; **SEX** 63% male
INTERVENTION	Both groups included exercise-based physiotherapy (specific versus control exercises) after an initial steroid injection; **DURATION** 12 weeks
PROGNOSTIC FACTORS CONSIDERED	**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)†
OUTCOME	Choice of surgery (yes/no, based on record of treatment)
ENDPOINT	After 1 year (after inclusion or after surgery)
SELECTION OF FACTORS	**FOR MULTIVARIABLE MODELLING** unclear (lack of information); there is no suggestion of predictor selection based on univariable analysis. **WITHIN MULTIVARIABLE MODELLING** unclear (lack of information). No rationale was provided for the combinations of prognostic factors, and no ‘final’ model was specified, but apparently, no stepwise regression was used.
STATISTICAL ANALYSIS	Logistic regression. Four models were presented‡.
MOST COMPLETE MODEL INCLUDING MAIN EFFECTS FOR ALL PROGNOSTIC MODELS	N outcome events = 41 Pseudo R^2: 0.28

Predictor/statistics§	OR	95% CI
Intact cuff	1.00	
PTT	0.92	(0.24; 3.46)
FTT	2.88	(0.32; 25.59)
Control vs. specific	8.68	(2.75; 27.37)
CM 1. quartile	1.00	
CM 2. quartile	0.42	(0.10; 1.82)
CM 3. quartile	0.11	(0.03; 0.47)
CM 4. quartile	0.12	(0.03; 0.58)
Calcification	2.59	(0.68; 9.85)
Degeneration	2.05	(0.43; 9.71)
Women	0.32	(0.09; 1.12)
Regression constant	0.01	(0.00; 0.46)

	FURTHER EVALUATION OF MODEL PERFORMANCE (including internal and external validation) None presented
PROGNOSTIC INDEX/STATEMENT	None presented

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

	Hung 2010
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) **PHASE OF RESEARCH** development
SETTING	Taiwan; orthopaedic **[**presumably outpatient] clinic in a national university hospital **STUDY DATES** unspecified
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 R^2: 0.73
	FLEX-SF score — cut-off < 41#
	Scapular internal rotation at 30° shoulder elevation (descending phase, unloaded) — cut-off < 0.7#
	Serratus anterior force as % of body weight — cut-off < 27%#
	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	‖All potential prognostic factors were dichotomised using cut-points derived from ROC analyses. ¶Apparently an adjustment variable. #Resulting values from sensitivity and specificity ROC analysis.

	Kromer 2014
OBJECTIVE	To analyze "…to what degree fear-avoidance beliefs and catastrophizing contribute to the variance of disability at baseline and at 3-month follow-up in patients with SPS [subacromial pain syndrome]."
DESIGN	Cohort study derived from 2 group RCT; consecutive recruitment **PHASE OF RESEARCH** development
SETTING	Germany; outpatient physiotherapy practices **STUDY DATES:** recruitment took place over a 18-month period; dates are unspecified
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 **MEAN AGE** 51.8 years; **SEX** 50% male
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 (physiotherapy for 5 weeks + continuation of home exercises for 7 weeks)
PROGNOSTIC FACTORS CONSIDERED	N ≥ 7**: Age††, 11-point VNRS††, FABQ-PA, PCS, Sex††, SPADI-F, symptom duration††
OUTCOME	SPADI-F change score
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	Hierarchical linear regression. The seven predictors were categorized into: demographic, clinical and pyschological factors.
FINAL MODEL	N outcome events: 88 R^2: .48; R^2 adjusted: .44

Predictor/statistics	Beta*	95% CI‡‡
Age	0.000	-0.32; 0.32
Sex	0.081	*-3.94, 9.86*
Duration of complaints	-0.324	*-0.06, -0.02*
SPADI-F baseline score	0.600	0.40, 0.78
FABQ-PA score	-0.102	*-1.14, -0.36*
PCS score	0.083	-0.23, 0.59

	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:

	• Conservative group: 11.3 (1.8) • Surgical group: 16.1 (1.7)
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, expecially 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 R^2 adjusted: .68 Acromial morphology Duration of symptoms — SEE = 0.76 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
PROGNOSTIC INDEX/STATEMENT	None presented
CALIBRATION, DISCRIMINATION, VALIDATION	None presented
PROGNOSTIC INDEX/STATEMENT	None presented
STUDY AUTHORS' CONCLUSIONS	"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."
NOTES	||||As reported by the authors, but there is a discrepancy. Of 128 eligible patients, 93 consented & 13 were excluded from the analysis, giving a sample of 80. ¶¶Erroneously analysed as continuous data in the regression.

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[59])