Shoulder Pain

Kuijpers, A.C.

2005

This study was supported by a grant (No. 904-65-09) from the Netherlands Organisation for Scientific Research (NWO), The Hague, The Netherlands.

ISBN: 90-6464-253-2

Cover design&Lay-out, One Fish/Gouw#

Printed by, Ponsen&Looijen bv, Wageningen, The Netherlands

VRIJE UNIVERSITEIT

Shoulder Pain

Prediction of outcome in primary care

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. T. Sminia,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Geneeskunde
op vrijdag 14 oktober 2005 om 10.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Antonius Christianus Kuijpers

geboren te Oisterwijk

promotor:	prof.dr. L.M. Bouter
copromotoren:	dr. D.A.W.M. van der Windt
	dr. G.J.M.G. van der Heijden

Contents

	Voorwoord	7
1	Introduction	11
2	Systematic review of prognostic cohort studies on shoulder disorders	19
3	Costs of shoulder pain in primary care	47
4	Clinical prediction rules for the prognosis of shoulder pain in general practice	63
5	Generalisability of clinical prediction rules for the prognosis of shoulder pain in general practice	89
6	A prediction rule for shoulder pain related sick leave	107
7	Generalisability of a prediction rule for shoulder pain related sick leave	127
8	Are psychological factors associated with persistent pain and disability in patients with shoulder pain or low back pain?	143
9	General discussion	163
	Summary	181
	Samenvatting	189
	Thanks to	195
	Curriculum Vitae	197

Voorwoord

Tijdens mijn promotietraject ben ik voortdurend in goed gezelschap geweest. Met goed gezelschap bedoel ik mensen met tegenwoordigheid van geest, uitstekend methodologisch inzicht en prettig in de omgang. De volgende mensen hebben in belangrijke mate bijgedragen aan de totstandkoming van dit proefschrift: Daniëlle van der Windt, Geert van der Heijden, Lex Bouter, Joan Boeke, Maurits van Tulder, Jos Twisk, Yvonne Vergouwe, Gert Bergman, Jacques Geraets, Camiel de Bruijn, en Marjan Uittenbosch.

Met stip op één staat Daniëlle, mijn co-promotor en dagelijkse begeleider, door wie mijn promotie een prettige, kalme en ontspannen tijd is geworden. Daan, je begon vriendelijk door me 3 maanden later te laten beginnen dan gepland en me de kans te geven door een deel van Azië te trekken. Dat was pas het begin. Je groeide door naar aardig toen ik ook nog eens een flinke reis naar Nieuw-Zeeland kon maken. Je piekte toen ik eind 2003 zonder problemen 2 maanden naar Midden-Amerika kon. Hier ben ik je erg dankbaar voor. Je gebrek aan methodologisch inzicht, (Engelse) schrijfvaardigheid, organisatorisch talent en toegankelijkheid zou hier ruimschoots mee gecompenseerd kunnen worden. Van een gebrek is echter geen sprake. In tegendeel zelfs! De combinatie van deze eigenschappen maakt je een begeleider die iedereen zich wenst en waar menigeen jaloers op is. Voor alles wat ik van je geleerd heb, dat je me mijn gang hebt laten gaan en me niet te veel achter mijn broek hebt aangezeten: Bedankt!

Geert, co-promotor op afstand. Eerst in Maastricht later in Utrecht. In het bezit van een uitmuntend methodologisch inzicht en het vermogen methodologische valkuilen te signaleren waarvan anderen nog geen flauw idee hebben. Altijd bereid mee te denken, bijvoorbeeld over nieuwe banen! Bedankt voor de vruchtbare samenwerking.

Lex, promotor. In het bezit van bovenmenselijke methodologische en didactische kwaliteiten. Ik heb genoten van jouw 'Lexiaans': heldere, klare, sturende taal met een vleugje ironie.

Joan, je beloofde 25 patiënten en het werden er 47, briljant! Jos, garantie voor een solide statistische basis van dit proefschrift! Helderheid boven alles. Maurits, met jou samenwerken is erg relaxed en de kostenstudie was inderdaad een appeltje-eitje. Yvonne, dank voor het wijzen van de weg in de wereld van predictieregels. Ik heb enorm veel van je geleerd! Gert, Jacques en Camiel, dank voor het plezier, de erg prettige samenwerking en het verzamelen van de data in de Groningse en Maastrichtse takken van het onderzoek. Marjan, door jouw kordate en doelgerichte optreden werd de Amsterdamse dataverzameling een ongekend succes!

Het was een eer met jullie te mogen samenwerken en dankzij jullie ligt er een helder en samenhangend proefschrift.

Ton Kuijpers
Amsterdam, 27 juni 2005

1

Introduction

Definition

In this thesis shoulder pain is defined according to the 1999 version of the Dutch guidelines for shoulder complaints, issued by the Dutch College of General Practitioners.[1;2] Shoulder pain is characterised as pain in the deltoid and upper arm region, as illustrated in Figure 1. Pain and stiffness are the prominent complaints. Pain and stiffness restrict the use of the arm and therefore limit daily activities, especially when using the hands above shoulder level. Lying on the affected shoulder is painful, which means that severe shoulder pain can cause problems with sleeping. Shoulder pain is associated with increased sick leave and incapability of performing daily activities.[3]

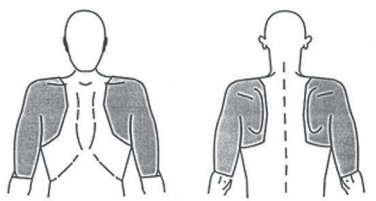


Figure 1 Location of shoulder pain

Prevalence / incidence

Shoulder pain is common with a one year prevalence ranging between 5% and 47%.[3-7] The prevalence in the general population in The Netherlands has recently been estimated at 17%.[8] The annual incidence of consultation for a new episode of shoulder pain in Dutch general practice ranges between 12 and 25/1000/year.[5;8-10]

Course and prognosis

Shoulder pain has an unfavourable outcome in many patients. Only about 50% of all new episodes of shoulder pain presented in primary care show complete recovery within 6 months[11-13], after 1 year this proportion

increases to only 60%.[12] In a narrative review Van der Heijden (1999) concluded that evidence for all reported prognostic factors was weak, and that most studies appeared to be of relatively poor methodological quality. Little was known about the prognostic value of psychosocial factors.[3]

The Dutch Shoulder Study

The 2nd edition of the Dutch guidelines for shoulder pain, issued by the Dutch College of General Practitioners[1;2], was published in 1999. Regarding several topics there was lack of evidence to provide clinicians with evidence based recommendations. Information about the predictive value of potential prognostic factors, above all occupational and psychosocial factors, was lacking. There was insufficient evidence for several interventions such as manual therapy, (graded) exercise therapy, and cognitive behavioural interventions. Besides, there was hardly any evidence on the costs associated with shoulder pain. This lack of information resulted in the development of a large research program seeking valid and applicable evidence on shoulder pain in primary care: The Dutch Shoulder Study (DSS).[14]

The DSS is a comprehensive cohort study, carried out between January 2000 and May 2005. The DSS consists of a prognostic cohort study (presented in this thesis) and three randomised controlled trials, which were carried out alongside each other. Between January 2001 and June 2003, 103 general practitioners (GP) recruited patients at first consultation for a new episode of shoulder complaints in three geographic areas in the Netherlands (Amsterdam, Groningen and Maastricht). All patients in the DSS had to meet the same general inclusion criteria, and specific additional inclusion criteria if eligible for a trial. Patients not eligible for a trial were invited for participation in the prognostic cohort study. The Groningen Manipulation Study (GMO)[15;16] studies the effectiveness of manipulative therapy for the shoulder girdle in addition to usual GP care. In two other trials a Graded Exercise Therapy (GET)[17] and an Education and Activation Program (EAP)[18] were studied. Baseline and follow-up assessments for all patients in the core dataset of the DSS were identical.

Clinical prediction rules

Knowing more about the prognostic value of clinical, psychosocial, and occupational factors in patients with shoulder pain will help to provide patients with adequate information regarding their risk of persistent symptoms on the short term (6 weeks) and long term (6 months). Such information may also support decisions regarding treatment and referral of patients. In our cohort study we developed clinical prediction rules to provide physicians with easy-to-use tools for calculating an individual's risk of persistent shoulder symptoms or shoulder pain related sick leave on the short term and long term.

Objectives

The objectives of this thesis originate from discussions in the editorial committee of the 2nd edition of the Dutch guidelines for shoulder pain of the Dutch College of General Practitioners.[1;2] The overall aim of this thesis was to provide the GP and patients with better information on the course and prognosis of shoulder pain. The objectives were:

1. To systematically review the available literature regarding prognostic factors in patients with shoulder pain.
2. To determine the costs of shoulder pain during 6 months following first consultation in general practice.
3. To develop clinical prediction rules for calculating the risk of persistent shoulder symptoms for individual patients, at 6 weeks and 6 months after first consultation in general practice.
4. To develop a clinical prediction rule for calculating the risk of shoulder pain related sick leave for individual workers, during 6 months following first consultation in general practice.
5. To evaluate the generalisability of the prediction rules to other populations of patients with a new episode of shoulder pain consulting their general practitioner.
6. To study the effect of psychological factors on the risk of a poor outcome in patients consulting a general practitioner for a new episode of shoulder pain.

Outline of this thesis

Chapter 2 describes the results of a systematic review in which the available evidence from 16 studies regarding predictors of outcome of shoulder disorders is summarized.[19] In chapter 3 we summarize the costs associated with shoulder pain in the 6 months following first consultation in primary care. We present direct health care costs as well as indirect costs related to loss of productivity. In chapter 4 we present the design and results of our prognostic cohort study among patients with shoulder pain in general practice. We describe the development of prediction rules for calculating the risk of persistent shoulder symptoms for individual patients, at 6 weeks and 6 months after first consultation in general practice. In chapter 5 the generalisability of these prediction rules is tested within the framework of the DSS. The prediction rules derived from the prognostic cohort study, are tested on the merged control groups of the three trials embedded in the DSS. In chapter 6 we describe the development of a prediction rule for shoulder pain related sick leave in the 6 months following first consultation in general practice. For this study we selected those patients from the cohort who reported paid work at enrolment. Chapter 7 describes the generalisability of the prediction rule for shoulder pain related sick leave. Again, the merged control groups of the embedded trials within the frame work of the DSS were used as a validation cohort. Furthermore, the generalisability of the prediction rule for sick leave was tested in another cohort of patients with shoulder pain, in which largely similar data were collected.[20] Chapter 8 compares the effects of psychological factors on the risk of a poor outcome in patients consulting a general practitioner for a new episode of shoulder pain or low back pain. This study was carried out to investigate whether similar mechanisms play a role in the transition from acute to chronic pain in patients with different types of pain. Finally, in chapter 9 the results of this thesis are critically reviewed, and recommendations for future research are given.

References

1. The Dutch College of General Practitioners (NHG) Practice Guideline. Accessed at http://nhg.artsennet.nl/upload/104/guidelines2/E08.htm on 1 september 2003.
2. Winters JC, Jongh AC, van der Windt DA, Jonquiere M, Winter AF, van der Heijden GJ et al. Practice Guideline "Shoulder complaints' [In Dutch: NHG-Standaard Schouderklachten]. Huisarts en wetenschap. 1999;42:222-31.
3. Van der Heijden GJMG. Shoulder disorders: a state-of-the-art review. Baillieres Best Pract Res Clin Rheumatol. 1999;13:287-309.
4. Luime JJ, Koes BW, Hendriksen IJ, et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. Scand J Rheumatol 2004;33:73-81.
5. Bot SD, Van der Waal JM, Terwee CB, Van der Windt DAWM, Schellevis FG, Bouter LM et al. Incidence and prevalence of complaints of the neck and upper extremity in general practice. Ann Rheum Dis 2005;64:118-23.
6. Bongers PM. The cost of shoulder pain at work. BMJ. 2001;322:64-65.
7. Picavet HS, Schouten JS. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3)-study. Pain. 2003;102:167-78.
8. Van der Linden MW, Westert GP, De Bakker DH, FG Schellevis. Complaints and disorders in the population and general practice [In Dutch: Klachten en aandoeningen in de bevolking en in de huisartsenpraktijk]. 2005 Utrecht/Bilthoven: NIVEL/RIVM.
9. Van der Windt DAWM, Koes BWK, Jong BAD, Bouter LM. Shoulder disorders in General Practice: incidence, patients characteristics, and management. Annals of Rheumatic Diseases. 1995;54:959-64.
10. Okkes IM, Oskam SK, Lamberts H. Van klacht naar diagnose: Episodegegevens uit de huisartspraktijk. Bussum. Coutinho, 1998.
11. Croft P, Pope D, Silman A. The clinical course of shoulder pain: prospective cohort study in primary care. BMJ 1996;313:601-2.
12. Van der Windt DAWM, Koes BWK, Boeke AJP, et al. Shoulder disorders in general practice: prognostic indicators of outcome. Br J Gen Pract 1996;46:519-23.
13. Winters JC, Sobel JS, Groenier KH, et al. The long-term course of shoulder complaints: a prospective study in general practice. Rheumatology 1999;38:160-3.
14. Van der Heijden GJMG. A comprehensive prognostic cohort study on shoulder disorders, with randomised controlled trials in subcohorts (Grant No. 904-65-09). Netherlands Organization for scientific Research (NWO), The Hague, The Netherlands 1999.
15. Bergman GJ, Winters JC, Van der Heijden GJMG, Postema K, Meyboom-de Jong B. Groningen Manipulation Study. The effect of manipulation of the structures of the shoulder girdle as additional treatment for symptom relief and for prevention of chronicity or recurrence of shoulder symptoms. Design of a randomized controlled trial within a comprehensive prognostic cohort study. J Manipulative Physiol Ther 2002;25:543-9.
16. Bergman GJ, Winters JC, Groenier KH, Pool JJ, Meyboom-de Jong B, Postema K et al. Manipulative therapy in addition to usual medical care for patients with shoulder dysfunction and pain: a randomized, controlled trial. Ann Intern Med 2004;141:432-9.

17. Geraets JJ, Goossens ME, De Bruijn CP, Koke AJ, De Bie RA, Pelt RA et al. A behavioural treatment for chronic shoulder complaints: concepts, development, and study design. Aust J Physiother 2004;50:33-8.

18. De Bruijn CP, De Bie R, Geraets JJ, Goossens ME, Koke AL, Van den Heuvel HW et al. Evaluation of an education and activation programme to prevent chronic shoulder complaints: design of an RCT [ISRCTN71777817]. BMC Fam Pract 2005;6:7.

19. Kuijpers T, Van der Windt DAWM, Van der Heijden GJGM, Bouter LM. Systematic review of prognostic cohort studies on shoulder disorders. Pain. 2004;109:420-431.

20. Bot SD, Van der Waal JM, Terwee CB, Van der Windt DAWM, Scholten RJPM, Bouter LM, et al. Predictors of outcome in neck and shoulder complaints: a cohort study in general practice. Spine (in press) 2005.

2

Systematic review of prognostic cohort studies on shoulder disorders

Ton Kuijpers
Daniëlle AWM van der Windt
Geert JMG van der Heijden
Lex M Bouter

Pain 2004;109:420-43

ABSTRACT

Background Shoulder complaints are common and have an unfavourable outcome in many patients. Only 50% of all new episodes of shoulder disorders end in complete recovery within 6 months. There is no consensus about prognostic indicators that can identify patients at high and low risk of chronicity.

Methods By a systematic search of the literature we identified 16 studies focusing on the prognosis of shoulder disorders.

Results The methodological quality of these 16 studies was assessed. Six of these were considered to be of relatively 'high quality'. There was a wide variety among the studies in length of follow-up, study population, evaluated prognostic factors, type of outcome measure and method of analysis. Due to this large heterogeneity, we refrained from statistical pooling. Instead, we used a best-evidence synthesis. There is strong evidence that high pain intensity predicts a poorer outcome in primary care populations and that middle age (45-54) is associated with poor outcome in occupational populations. There is moderate evidence that a long duration of complaints, and high disability score at baseline predict a poorer outcome in primary care.

These results need to be interpreted with caution because of the small number of studies on which these conclusions are based, and the large heterogeneity among studies regarding follow-up, outcome measures, and analysis.

Introduction

Shoulder disorders are common. The one year prevalence in various studies ranges between 5% and 47%.[1;2] The point prevalence in the general population in The Netherlands has recently been estimated at 21%.[3] In a British study a lower point prevalence of 14% has been found.[4] The annual incidence of shoulder disorders in Dutch general practice ranges between 12 and 25/1000/year.[5;6]

Shoulder complaints have an unfavourable outcome in many patients. Only about 50% of all new episodes of shoulder complaints presented in primary care show complete recovery within 6 months[7-9], after 1 year this proportion increases to only 60%.[9]

Van der Heijden reviewed the literature on prognostic indicators of a favourable outcome within 3 months, in a narrative way. Evidence for all reported factors was weak, and most studies appeared to be of relatively poor methodological quality. Little is known about the prognostic value of psychosocial factors. It is suggested that psychosocial factors such as inadequate pain cognitions and pain behaviour, are likely to predict a poor outcome of painful musculoskeletal conditions.[2]

It is of importance for clinical practice to know more about the prognostic value of clinical, psychosocial, and occupational factors in patients with shoulder disorders. It may help to provide patients with adequate information regarding the most likely course of their symptoms. Health care providers need prognostic information to distinguish between patients with a favourable outcome and those with a high risk of chronic shoulder pain and disability. This may facilitate decisions regarding treatment and referral of patients. However, no attempts have been made to conduct a systematic search of the literature and to summarise the available evidence regarding prognostic factors of shoulder disorders.

Methods

Identification and selection of the literature

We conducted a systematic, computerised search of the literature based on recommendations by Haynes et al.[10] in Medline (1966 through February

2003), Embase (1991 through February 2003), Cinahl (1982 through February 2003), Psychinfo (1967 through February 2003), Sportdiscus (1949 through February 2003). The following key words and medical subject headings were used: shoulder, shoulder pain, shoulder joint, shoulder injuries, shoulder impingement syndrome, prognos*(truncated), predict*(truncated), course, clinical study, longitudinal study, prospective study and retrospective study. The citations we found were screened by two reviewers independently (TK and DW). The reference lists of all selected publications were checked to retrieve relevant publications which had not been found with the computerised search. The publications had to meet the following selection criteria:

- the study focussed on patients suffering from shoulder complaints
- the association (ORs or RRs, with corresponding p-value or 95% CI) of at least one prognostic factor with the outcome of shoulder pain had to be presented
- the design had to be a cohort study
- the article was published in English
- results were published as a full report before February 2003
- studies that focussed on shoulder pain due to luxation, cancer or systemic diseases such as rheumatoid arthritis or osteoporosis were excluded. Also studies that focussed on the results of surgery were excluded.

Quality assessment

The methodological quality of each of the studies was assessed independently by three reviewers (TK, DW, and GH). A standardised checklist of predefined criteria was used, which is a modified version of the checklists by Scholten et al.[11], Borghouts et al.[12], and Hudak et al.[13], and is based on theoretical considerations and methodological aspects described by Altman[14] and Hudak et al.[13] (Table 1). Disagreements among the reviewers were discussed during a consensus meeting. In case of persisting disagreement between 2 reviewers it was the third that made the final

decision. The checklist covers aspects of internal validity (criteria A, D, E, F, G, H, I, J, K, L, M, P, Q), generalisability (criteria B, C, N, O) and precision (criterion R), which are all of great importance in descriptive epidemiological studies[14]. The list contains 7 categories: study population, response rate, follow-up, treatment, outcome, prognostic factors and data presentation. The list contains 18 criteria which can be scored positive ('+'), negative ('-') or 'unclear ('?'). If an item is scored as 'unclear' it means that the paper provides insufficient information about this criterion. A positive score indicates sufficient information and a positive assessment. A negative score indicates sufficient information, but potential bias due to inadequate design or conduct. A negative score can only be assigned to criteria of internal validity. Exceptions are criteria N, O and R, because an 'unclear' did not make sense here. A more detailed explanation of each criterion is given in the Appendix.

The maximum attainable score on the criteria list is 18 points. The total score is the sum of all the criteria which are scored positive, negative scores are not subtracted. A priori, we chose to consider a study of 'high quality' when it scores more than 10 points (>60% of the maximum attainable score) and of 'low quality' when it scores ≤10 points. Sensitivity analyses were conducted to assess the robustness of this cut-off point, that is, whether this change will lead to different conclusions.

Data extraction

Data were extracted of the selected studies regarding study population, design, setting, outcome measures, prognostic factors and strength of association. To facilitate interpretation and comparison of the results the studies are categorised per setting (primary care, secondary care and occupational setting). When not given, and sufficient data were available, for each study the univariate association was calculated between prognostic factors and outcome in terms of Risks Ratios (RR) or Odds Ratios (OR) with 95% confidence intervals (CI). Univariate, or if available multivariate associations were presented in tables.

Table 1 Criteria list for assessing the methodological quality of prognostic cohort studies on shoulder pain

Criteria		Score
Study population		
A.	Inception cohort (defined in relationship to onset of symptoms)	+ / - / ?
B.	Description of inclusion and exclusion criteria	+ / ?
C.	Description of studypopulation	+ / ?
Response		
D.	Response ≥ 75%	+ / - / ?
E.	Information about non-responders versus responders	+ / - / ?
Follow-up (extent and length)		
F.	Prospective data collection	+ / - / ?
G.	Follow-up of at least 6 months	+ / - / ?
H.	Drop-outs/loss to follow-up < 20%	+ / - / ?
I.	Information completers versus loss to follow-up/drop-outs	+ / - / ?
Treatment		
J.	Treatment in cohort is fully described/standardised	+ / - / ?
Outcome		
K.	Standardised assessment of relevant outcome criteria	+ / ?
Prognostic factors		
L.	Standardised assessment of patient characteristics and potential clinical prognostic factor(s)	+ / ?
M.	Standardised assessment of potential psychosocial prognostic factor(s)	+ / ?
Data presentation		
N.	Frequencies of most important outcome measures	+ / -
O.	Frequencies of most important prognostic factors	+ / -
P.	Appropriate analysis techniques	+ / - / ?
Q.	Prognostic model is presented	+ / - / ?
R.	Sufficient numbers	+ / -

+, positive (sufficient information and a positive assessment); - (sufficient information, but potential bias due to inadequate design or conduct); ?, unclear (insufficient information)

Analysis

Depending on homogeneity in study population, type of prognostic factors, outcome measures, and study quality, statistical pooling was considered. When a pooled estimate could not be computed, a qualitative analysis (best evidence synthesis) was performed to summarize the value of the prognostic indicators. In this analysis the available evidence for a prognostic factor was summarised by taking into account the number of studies evaluating this factor, the methodological quality of these studies,

and the consistency of the available evidence. We present prognostic factors which showed in at least one study a RR or OR above 2.0 or below 0.5, or a statistically significant ($p<0.05$) association. We did not want to depend solely on statistical significance, as many cohorts included in the review were rather small, and relevant associations between prognostic factors and outcomes may have remained undetected. Findings were considered consistent if ≥75% of the studies which reported a factor showed the same direction of the association. In Table 2 we defined four levels of evidence which are based on Sackett et al.[15] and Ariëns et al.[16] (Table 2).

Table 2 Levels of evidence for prognostic factors on shoulder disorders

Level of evidence	
Strong	Consistent findings (≥75%) in at least 2 high quality cohorts
Moderate	Consistent findings (≥75%) in one high quality cohort and at least one low quality cohort
Weak	Findings of one high quality cohort or consistent findings (≥75%) in at least 3 or more low quality cohorts
Inconclusive	Inconsistent findings irrespective of study quality, or less than 3 low quality cohorts available
No evidence	No data presented

Results

Selection of studies

We found 1273 citations (468 Pubmed, 507 Embase, 211 Cinahl, 54 Psychinfo, 33 Sportdiscus). Out of this number 48 abstracts seemed to fulfil the selection criteria and the full publications of these were retrieved. When assessing the full publications some turned out to focus on rotator cuff tears (n=2), some papers aimed at aetiology instead of prognosis (n=4), some dealt with treatment (n=7), and some with diagnoses (n=1). Not presenting a separate analysis for shoulder disorders (n=18) was a major reason for excluding papers. Finally, 16 papers were included and the methodological score was assessed.

Table 3 Results of the methodological assessment of prognostic cohort studies on shoulder disorders

First author	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	Quality Score*	Score (%)
Cassou[24]	+	+	+	+	?	+	+	+	?	?	+	+	+	+	+	+	+	+	15	83
Brox[19]	-	+	+	-	?	+	+	+	?	+	+	+	+	+	+	+	+	+	14	78
Van der Windt[9]	+	+	+	+	-	+	+	+	-	-	+	+	?	+	+	+	+	+	14	78
Macfarlane[25]	-	?	+	+	?	+	+	-	+	?	+	+	+	+	+	+	+	+	13	72
Miranda[17]	?	?	+	+	+	+	+	?	?	?	+	+	+	+	-	+	+	+	12	67
Chard[30]	+	?	+	?	?	+	+	+	?	?	+	+	?	+	+	+	?	+	11	61
Morrison[20]	?	+	-	?	?	+	+	+	?	+	+	+	?	+	+	-	-	+	10	56
Bartolozzi[21]	-	+	-	?	?	+	+	-	?	+	+	+	?	+	+	+	-	+	10	56
Viikari-Juntura[18]	+	?	+	-	+	+	-	?	?	?	+	?	?	+	+	+	+	+	10	56
Binder[22]	-	+	-	?	?	+	+	+	?	+	+	+	?	+	+	-	-	-	9	50
Solomon[26]	-	+	+	?	?	+	+	?	?	?	+	+	?	-	+	+	+	-	9	50
Croft[7]	+	?	+	?	?	+	+	-	+	?	+	?	?	+	-	+	?	+	9	50
Shaffer[27]	-	+	+	?	?	+	+	-	?	-	+	?	?	+	-	-	-	-	6	33
Kaergaard[28]	?	?	+	+	?	+	+	-	-	?	+	?	+	-	-	?	?	-	6	33
Kuroda[29]	?	?	+	?	?	+	+	?	?	?	?	?	?	+	-	?	?	+	5	28
Mulcahy[23]	-	?	+	?	?	+	-	-	?	+	+	?	?	?	-	-	-	-	4	22

*(total '+')

Methodological quality

The results of the quality assessment are presented in Table 3. The overall quality score ranged from 4 to 15 points, with a median score of 10 points. Using our cut-off point of >10 points, six studies were classified as high quality studies. The items of the criteria list which most often (>8 of the 16 publications) obtained a negative score were 'Inception cohort' (item A), 'Adequate response rate' (item D), 'Information about responder/ non-responders' (item E), 'Follow-up >6 months' (item H), 'Adequate information about loss to follow-up' (item I), 'Treatment described/ standardised' (item J), 'Assessment of psychosocial factors' (item M) and 'Prognostic model presented' (item Q). Only 2 studies[17;18] presented information about response-rate and information about characteristics of responders versus non-responders in order to evaluate whether the response was selective or not. Only 5 studies[19-23] presented information regarding treatment and whether it was standardized. Seven studies[9;17-19;24-26] used adequate methods to compose a multivariable prognostic model. The studies[23;27-29] with a method score in the lowest tertile of the scale (≤33%) all suffered from inadequate data presentation (item N, O, P, Q, R).

Study characteristics

Table 4 (see end of chapter) summarises the study characteristics of the publications including study population, outcome measures, follow-up, prognostic factors and the strength of the association with their 95% confidence interval. Four studies were conducted in a primary care setting, another 4 in an occupational setting and 8 in a hospital setting. Most frequently reported prognostic factors were pain, duration of complaints, age and gender. A few studies[17;19;24] assessed the value of psychosocial predictors. There was considerable variation among the studies with respect to the length of follow-up (range 2 months to 7 years), type of outcome measure (pain, disability, recovery, sick leave, ROM, different shoulder questionnaires) and method of analysis (univariable vs. multivariable). Hence, we considered statistical pooling to be not sensible, and therefore used a best-evidence synthesis to summarize the importance of prognostic factors (Table 2).

Levels of evidence

In Table 5 we only present those prognostic factors which in at least 1 study showed RR or OR above 2.0 or below 0.5 or a statistically significant ($p<0.05$) association. Most factors were only measured in one study, and consequently their prognostic value remains uncertain. There is, however, strong evidence that high pain intensity predicts a poorer outcome[9;25] in primary care populations and that middle age (45-54)[17;24] is associated with poor outcome in occupational populations (Table 5). In addition, there is moderate evidence that a long duration of complaints, and high disability score at baseline predict a poorer outcome in primary care (Table 5). Factors with RR or OR between 0.5 and 2.0 or a not statistically significant association were, for example, years of education, repetitive work, precipitating trauma and instability of the glenohumeral joint (Table 4).

Table 5 Overall level of evidence for prognostic factors and their association with (long term) poorer outcome

Prognostic factor	Outcome	QS>60%	QS≤60%	Level of evidence
Primary care				
Sick leave at baseline	Neer-score	1/1 (100%)	-	Weak
Regular medication	Neer-score	1/1 (100%)	-	Weak
Concomitant neck pain	Symptoms	1/1 (100%)	-	Weak
High pain intensity	Symptoms	2/2 (100%)	-	Strong
No precipitating trauma	Symptoms	1/1 (100%)	-	Weak
No acute bursitis	Symptoms	1/1 (100%)	-	Weak
Long duration of complaints	Disability, pain	1/1 (100%)	1/1 (100%)	Moderate
High disability score	Pain	1/1 (100%)	1/1 (100%)	Moderate
Previous episodes of pain	Pain	-	1/1 (100%)	Inconclusive
Severe restricted passive elevation (<101°)	Disability	-	1/1 (100%)	Inconclusive
Occupational setting				
Middle aged	Symptoms	2/2 (100%)*	-	Strong
Previous musculoskeletal disorders	Symptoms	1/1 (100%)	-	Weak
High job demand	Symptoms	1/1 (100%)	-	Weak
Overload at work	Symptoms	1/1 (100%)	-	Weak
No sporting activities	Symptoms	1/2 (50%)	0/1 (0%)	Inconclusive
Worker group (blue vs. white color)	Sick leave	-	1/1 (100%)	Inconclusive
Sick leave (preceding examination)	Sick leave	-	1/1 (100%)	Inconclusive
Duration of symptoms (0-2 vs. >7 days)	Sick leave	-	1/1 (100%)	Inconclusive
Continuous high intensity pain	Sick leave	-	1/1 (100%)	Inconclusive
Rotation of head (pain)	Sick leave	-	1/1 (100%)	Inconclusive
Abduction of arm (pain)	Sick leave	-	1/1 (100%)	Inconclusive

Table 5 Continued

Prognostic factor	Outcome	QS>60%	QS≤60%	Level of evidence
Secondary care				
Gradual onset	UCLA-score	-	1/3 (33%)[$]	Inconclusive
Long duration of complaints	UCLA-score	-	1/4 (25%)[$]	Inconclusive
Dominant side involved	ROM	-	1/4 (25%)	Inconclusive
Type acromion (type II or III)	UCLA-score	-	1/1 (100%)	Inconclusive
Tenderness acromion	UCLA-score	-	1/1 (100%)	Inconclusive
Severe functional impairment	UCLA-score	-	1/1 (100%)	Inconclusive
Weakness	UCLA-score	-	1/1 (100%)	Inconclusive
Moderate or large tear	UCLA-score	-	1/3 (0%)	Inconclusive
Manual work	ROM	-	1/1 (100%)	Inconclusive
Referral to specialist	Pain	-	1/1 (100%)	Inconclusive
Worse baseline pain	Pain	-	1/1 (100%)	Inconclusive
Worse baseline function	Function	-	1/1/ (100%)	Inconclusive
More education (per year)	Pain	-	1/1 (100%)	Inconclusive
More education (per year)	Function	-	1/1 (100%)	Inconclusive
Osteoarthritis	Pain	-	1/1 (100%)	Inconclusive
Continuing overhead sports	Symptoms	-	1/1 (100%)	Inconclusive
Age	Symptoms[#]	-	1/6 (100%)	Inconclusive

Only factors are presented which scored clinically relevant associations (RRs, ORs >2.0 or <0.5 or significant associations, p<0.05) in at least one study. QS, quality score; UCLA, Shoulder-Rating scale of the University of California at Los Angeles; ROM, range of motion; RR, relative risk; OR, odds ratio; * in 1 study[24] only significant for women; [#]2 studies reported age on ROM, n.s; [$]outcome is ROM

Psychosocial factors

There are a few studies[17;19;24;28;31] which considered psychosocial factors (locus of control, emotional distress, job demand, job control, mental stress). None of these studies showed RR or OR above 2.0 or below 0.5 or a statistically significant (p<0.05) association.

Discussion

The present paper is the first systematic review of the current literature on potential prognostic indicators of shoulder disorders. Van der Heijden[2] conducted a narrative review of the literature, and found the following prognostic indicators of a favourable outcome within 3 months: mild trauma preceding symptoms, early presentation, preceding overuse and heavy and unusual activities of the upper extremity, an acute onset, a high erythrocyte sedimentation rate, and restricted prescription and use of medication. Factors that were reported to predict a poor outcome at 3 months were severe pain at first presentation, a prior episode, a severe

restriction of the passive abduction range, diabetes mellitus, concomitant neck pain, cervical spondylosis and radicular symptoms, higher age, involvement of the dominant side and sick-leave from work. Evidence for each of these factors was weak, and most studies appeared to be of relatively poor methodological quality.

In our systematic review we found disappointingly little evidence for most factors which in current literature are suggested to be of prognostic importance. Caution is needed with the interpretation of the results of our analysis, because the majority of studies suffer from many flaws in the design and conduct. Yet, there is consistent evidence that high pain intensity in primary care populations and middle age (45-54) in occupational populations are strong predictors for a poor prognosis, while there is some evidence that long duration of complaints and high disability score at baseline are predictors for a poor prognosis in primary care populations. There were no studies of sufficient quality of methods in secondary care. To date, there is no evidence for the prognostic importance of psychosocial factors.

Only 16 studies met our inclusion criteria, of which 6 were of high quality of methods. Besides the overall lack of quality of methods there was considerable heterogeneity regarding design, study populations, prognostic factors and outcome measures. This heterogeneity impedes meta-analysis. Therefore we decided to perform a best evidence synthesis of the available evidence.

Limitations

We restricted our search to full papers published in English. However, the influence of language bias is disputed, and its effect has not been firmly established.[32-34] We searched in electronic databases that are considered to be important and relevant for the topic of our review. Yet, we may have missed studies which are not included in these databases and which were not identified during our additional reference checking, for instance non-journal publications or unpublished cohort studies. The addition of non-journal publications has been shown to move the effect estimates towards a null result.[35;36] Given the fact that our review could not demonstrate

strong evidence for many relevant prognostic factors, we do not believe that inclusion of unpublished material or non-journal publications would strongly influence our conclusions regarding prognostic factors in shoulder pain.

Levels of evidence

Any system for defining levels of evidence is arbitrary. We chose a system that has been used in a systematic review on prognostic factors for whiplash related disorders.[11] We believe to have used a robust cut-off point to identify studies of high quality of methods, although any cut-off point is arbitrary. With a cut-off point of 50% (instead of 60%) there is also weak evidence for the prognostic importance of sick leave, duration of symptoms, continuous high pain intensity, rotation of head and abduction of arm in the occupational setting, and for acromion type III, tenderness acromion, severe functional impairment, weakness, moderate or large tear change in the hospital setting. In contrast, with a cut off point of 70% there is less strong evidence for the prognostic importance of middle age (45-54) in occupational setting, while there is no evidence left for overload at work.

Outcome assessment

As can be seen in Table 4 there is wide variation in the use of outcome measures between studies. Although most studies used a standardised assessment for at least one outcome measure, outcome measures used differed from a validated questionnaire to percentages patients reporting recovery or persistent pain. Only few studies[9;18] reported results for both within and after 6 months follow-up. This variation between studies makes it very difficult to pool results or to draw consistent and firm conclusions regarding the predictive value of any prognostic factor.

Psychosocial factors

It is suggested that there is a relationship between psychosocial factors such as depression, catastrophizing and kinesiophobia, and the persistence

or recurrence of chronic musculoskeletal pain.[2;37;38] For shoulder pain the importance of these factors, and their putative mechanism are not clear. Perhaps partly the same mechanism plays a role. There is a need for sound research regarding the prognostic importance of these psychosocial factors in patients with shoulder disorders.

Recommendations

Systematically reviewing prognostic studies is still in development and no validated or widely used criteria list is available. But this review unmistakably shows the need for well-conducted prospective cohort studies on putative prognostic factors of shoulder disorders. Moreover, because of the few small studies on which our conclusions are based, and the high heterogeneity among studies regarding follow-up, outcome measures, and analysis, we feel that the results of this review need to be interpreted with considerable caution.

In our opinion an appropriate prospective cohort study should fulfil all the criteria of our checklist (Table 1). Such future studies should focus on the predictive value of socio-demographic and clinical factors, but in particular on psychosocial factors, notably distress, fear and avoidance, kinesiophobia, coping-styles and job demand and control for shoulder disorders. New evidence on these putative prognostic predictors will enable better decisions on the choice of interventions. Outcomes estimates preferably are to be expressed as absolute risks, instead of RRs or ORs. A multivariable prognostic analysis may help to generate a prognostic index for differentiation between patients at high and low risk of persistent shoulder complaints. Such index needs to be validated both internally, i.e. with a split sample technique in the same population and externally, i.e. tested on another population.[39] Such an index should allow care providers easily to predict the likelihood of recovery in, for example, 6 months for any patient.

Table 4 Summary of study characteristics of prognostic cohort studies on shoulder disorders

First Author	Study Quality (%)	Study population	Outcome measures/ Duration follow-up	Prognostic factor(s)	Strength of association (95% CI)
Primary care / population-based cohorts					
Brox[19]	78	Patients with diagnosis of rotator tendinosis, referred by general practitioners. N = 125 (participants of a RCT comparing surgery, exercises, and placebo laser), drop-out 9%	Neer shoulder score (0-100) Succes: ≥ 80 points (6 months)	Not on sick leave Not on regular medication Active treatment (ref = not active) Years of education, overhead work activity, comorbidity, isometric strength endurance, locus of control beliefs, emotional distress	Multivariate analysis, adjusted for age, gender, symptom duration, baseline Neer score: OR = 4.4 (1.6-12.1) OR = 4.2 (1.5-11.1) OR = 4.8 (1.7-13.6) n.s.
Van der Windt[9]	78	Patients with a new episode of shoulder pain (not consulted their GP in the preceding year). N = 349, drop-out 13%	persistent symptoms (12 months)	Concomitant neck pain High pain intensity Precipitating trauma Diagnosis (acute bursitis) Age, gender, arm dominance,	Multivariate analysis: OR = 2.8 (1.7-4.6) OR = 2.0 (1.2-3.3) OR = 0.4 (0.2-0.9) OR = 0.4 (0.2-0.8) n.s.

Table 4 continued

First Author	Study Quality (%)	Study population	Outcome measures/ Duration follow-up	Prognostic factor(s)	Strength of association (95% CI)
Macfarlane[25]	72	Shoulder pain (current or in the preceding month): self-report questionnaire N = 135, drop-out 18%	% shoulder pain (3 years)	Pain at baseline Symptom duration (>1 year) Shoulder related disability (≥5 items on 22-item questionnaire) Age, sex, GP visit, area of pain, sudden onset, distress (GHQ), restricted ROM	Multivariate analysis, adjusted for age and sex: OR= 3.1 (1.1-8.2) OR= 2.9 (1.1-7.7) OR= 3.1 (0.9-11.0) n.s.
Croft[7]	50	Patients with a new episode of shoulder pain in general practice N = 166, drop-out 25%	Validated 22-item disability questionnaire (6 months)	Baseline disability score > 10 points Symptom duration (>1 month) Injection at baseline Previous episodes of shoulder pain Severely restricted passive elevation (<101°)	poorer outcome ($p<0.05$) (Beta's not presented)

Table 4 continued

First Author	Study Quality (%)	Study population	Outcome measures/ Duration follow-up	Prognostic factor(s)	Strength of association (95% CI)	
Occupational medicine						
Cassou[24]	83	Workers born in 1938, 1943, 1948 and 1953 with chronic neck-shoulder pain (> 6 months), random sample from occupational physicians' files. N = 1804 (in final analysis), drop-out 12.6%	% disappearance of pain (5 years)		Multivariate anlysis:	
				Year of birth (ref=1953)	Men	Women:
				1948	OR = 1.5 (0.9-2.5)	OR = 0.8 (0.5-1.2)
				1945	OR = 1.2 (0.8-1.9)	OR = 0.6 (0.4-0.8)
				1938	OR = 1.0 (0.6-1.5)	OR = 0.6 (0.4-0.9)
				Repetitive work (ref = never)		
				In 1990	univariate only	0.8 (0.5-1.3)
				Before 1990	univariate only	0.5 (0.3-0.7)
				High job demand	OR = 0.7 (0.5-0.9)	0.7 (0.6-0.9)
				Previous musculoskeletal disorders	OR = 0.4 (0.3-0.6)	0.6 (0.5-0.8)
				Sporting activities	OR=1.5 (1.1-2.1)	univariate only
				Precise movements, awkward work, repetitive work, job control, shift work	n.s.	n.s.
Miranda[17]	67	Employees of a forestry company in Finland reporting severe shoulder pain: > 30 days in the preceding 12 months. N = 419, drop-out 12%	% persistent severe pain (12 months)	Individual factors:	Multivariate OR	
				Age		
				<35	OR = 1.0	
				35-44	OR = 0.9 (0.3-2.6)	
				45-54	OR = 3.6 (1.3-10.2)	
				≥55	OR = 1.6 (0.5-4.8)	
				Sports activity added score		
				>156 vs <52	OR = 0.7 (0.4-1.3)	
				Gender (female)	OR = 0.7 (0.4-1.2)	

Table 4 continued

First Author	Study Quality (%)	Study population	Outcome measures/ Duration follow-up	Prognostic factor(s)	Strength of association (95% CI)
				Overload at work (definite vs none)	OR=3.8 (1.8-8.0)
				Other work load factors (e.g. working above shoulder level), mental stress, body mass index	n.s.
Viikari-Juntura[18]	56	Patients seeking medical advice for a new episode of neck-shoulder pain at an occupational health service. N = 474, drop-out 0%	Sick leave >3 days (60 days)		Multivariate analysis:
				Worker group (blue collar)	OR = 6.8 (2.1-22.4)
				Sick leave preceding examination	OR = 6.5 (2.1-20.4)
				Symptom duration (>7 days vs. 0-2 days)	OR = 0.1 (0.0-0.3)
				Continuous pain	OR = 1.7 (0.5-5.7)
				High pain intensity	OR = 1.1 (0.3-4.0)
				Interaction continuous pain x intensity	OR = 5.2 (1.0-28.1)
				Pain during rotation of the head	OR = 7.8 (3.0-20.1)
				Pain in shoulder in abduction of arm	OR = 5.9 (2.7-12.7)
				Other symptoms and signs	n.s.
Kaergaard[28]	33	Female sewing machine operators with neck-shoulder disorders. N = 40, drop-out 30%	% recovery (2 years),	Work exposure	n.s.
				Physical activity at leisure time	n.s.

Table 4 continued

First Author	Study Quality (%)	Study population	Outcome measures/ Duration follow-up	Prognostic factor(s)	Strength of association (95% CI)
Chard[30]	61	Patients with rotator cuff tendinitis ≥ 6 months after their first attendance in a shoulder clinic. N=137, drop-out 6%.	Shoulder pain resolved (mean 19 months)	Precipitating Cause Unknown Injury Employment Overuse/strain Occupation Housewife/retired Manual Non-manual	Univariate analysis: RR=1.0 RR=0.76 (0.4-1.42) RR=0.63 (0.29-1.4) RR=1.30 (0.82-2.06) RR=1.0 RR=0.88 (0.52-1.51) RR=1.09 (0.67-1.76)
Secondary care					
Morrison[20]	56	Patients diagnosed with subacromial impingement syndrome at center for sports medicine (historic cohort study). N=667, drop-out 8%.	Shoulder rating system University of California Los Angeles (≥28 points). (mean 27 months follow-up)	Female Dominance Dominant Non-dominant Bilateraal Type Acromion Type I Type II Type III Tenderness acromioclavicular joint Yes/No Age <20 21-40 41-60 >60	Univariate analysis: RR=0.97 (0.86-1.09) RR= 1.0 RR= 0.92 (0.81-1.05) RR=0.87 (0.67-1.13) RR=1.0 RR=0.74 (0.65-0.84) RR=0.7 (0.61-0.8) RR=0.83 (0.7-0.98) RR=1.0 RR=0.88 (0.7-1.11) RR=1.00 (0.8-1.25) RR=0.76 (0.57-1.01)

Table 4 continued

First Author	Study Quality (%)	Study population	Outcome measures/ Duration follow-up	Prognostic factor(s)	Strength of association (95% CI)
				Onset	
				Acute	RR= 1.0
				Non-acute	RR=0.81 (0.7-0.94)
				Chronic	RR=0.86 (0.75-0.99
Bartolozzi[21]	56	Patients attending orthopaedic department with impingement syndrome treated non-operatively. N=170, drop-out=20%.	Shoulder rating system University of California Los Angeles (≥29 points) (mean 20 months, > 6 months)		Univariate analysis:
				Female	$RR_{female\ /\ male}$ = 1.03 (0.81-1.31)
				Age	$RR_{40\text{-}60\ /\ <40y}$ = 1.10 (0.78-1.56)
					$RR_{>60\ /\ <40y}$ = 1.23 (0.89-1.71)
				Dominance	$RR_{non\text{-}dominant/dom}$ = 0.95 (0.74-1.24)
				Onset of symptoms	$RR_{insidious/\ acute}$ = 0.97 (0.75-1.26)
				Duration of pre-treatment symptoms*	$RR_{3\text{-}6\ mnths/<3\ mnths}$ = 0.90 (0.72-1.14)
					$RR_{>6\ mnths/<3\ mnths}$= 0.69 (0.52-0.91)
				Functional impairment*	$RR_{moderate\ /\ mild}$ =0.81 (0.60-1.09)
					$RR_{severe\ /\ mild}$ = 0.65 (0.46-0.91)
				Recreational or occupational demands	$RR_{moderate\ /\ low}$ = 1.18 (0.85-1.64)
					$RR_{severe\ /\ low}$ = 1.25 (0.97-1.62)
				Instability	$RR_{present/absent}$ = 0.97 (0.64-1.46)
				ROM	$RR_{mild\ /\ none}$ = 0.85 (0.60-1.20)
					$RR_{moderate\ /\ none}$ = 0.78 (0.45-1.36)
					$RR_{severe\ /\ none}$ = 0.96 (0.63-1.46)
				Weakness	$RR_{yes/no}$ = 0.68 (0.49-0.93)
				Rotator cuff pathology*	
				Impingement or tendinitis	RR= 1.0
				Partial or small full thickness tear	RR= 0.82 (0.53-1.26)
				Moderate or large tear	RR= 0.34 (0.14-0.80)
				Treatment	No significant differences

Table 4 continued

First Author	Study Quality (%)	Study population	Outcome measures/ Duration follow-up	Prognostic factor(s)	Strength of association (95% CI)	
Binder[22]	50	Patients attending a rheumatology department with shoulder pain for > 1 month (diagnosis frozen shoulder). Participants in a RCT. N = 42, drop-out 5%	ROM (mean 44 months)	Non-dominant side involved	mean difference abduction better: 12° (p < 0.05)	
				Manual work	worse: 15° (p < 0.05)	
				Therapy: mobilisation versus injections, ice, or no additional treatment	worse: 15° (p < 0.05)	
				Age, sex, symptom duration	n.s.	
Solomon[26]	50	Consecutive patients presenting acute shoulder pain to general internists, rheumatologists, or orthopaedic surgeons. N=63, drop-out 21%	Pain and function: Shoulder Pain and Disability Index (12 months)	Multivariate assoc.:	Improvement pain	Improvement function
				Referred	β= -2.4, p=0.02	β= -1.4, p=0.17
				Worse baseline pain, per point	β= 4.2, p=0.0002	Not in the model
				Worse baseline function per point	Not in the model	β= 4.9, p=0.0001
				Older age per year	β= 1.2, p=0.24	β= -0.8, p= 0.46
				Female	β= 1.4, p=0.2	β= 1.3, p=0.20
				More education, per year	β= 3.3, p=0.0019	β= 2.1, p=0.047
				Longer pain duration per month	β= -2.2, p=0.038	β= -1.2, p=0.24
				Osteoarthritis	β= 2.3, p= 0.026	β= 1.0, p=0.34
				Rotator cuff tear	β= 1.4, p=0.19	β= 1.1, p=0.26

Table 4 continued

First Author	Study Quality (%)	Study population	Outcome measures/ Duration follow-up	Prognostic factor(s)	Strength of association (95% CI)
Shaffer[27]	33	Patients with a diagnosis of either adhesive capsulitis or frozen shoulder in an orthopaedic clinic (retrospective study). N=92, drop out 33%	ROM (mean 7 years)	Age, Dominance, Side, Acute or gradual onset, Minor trauma or spontaneous onset, Duration of symptoms at baseline, Treatment, Response to treatment, Bilateral involvement, Associated medical problems	n.s.
Kuroda[29]	28	Patients who visited a Shoulder Disorder Clinic (Hospital) with atraumatic shoulder instability. N=341	Recovery (≥3 years)	Stopping sports Overhead Non-overhead Female Age	RR for recovery: RR=8.67 (2.7-27.1) RR=1.37 (0.55-3.43) RR=0.94 (0.56-1.58) p=0.01
Mulcahy[23]	22	Patients with frozen shoulder, referred for arthrographic examination N=51, drop-out 25%	Better, unchanged, worse (<6 months)	Tears (vs. no tears)	RR=0.77 (0.47-1.28)

RR, relative risk; OR, odds ratio; CI confidence interval; ROM, range of motion; n.s., not significant
*Independent predictors in a multivariate analysis (No frequencies)

Appendix Explanation of the criteria from Table 1

A. Positive if patients were identified at an early uniform point (inception cohort) in the course of their disease (first episode -with restriction to duration of symptoms- of shoulder pain in lifetime or first treated episode of shoulder pain).

B. Positive if criteria were formulated for at least: age, duration of symptoms, relevant co-morbity (i.e cervical radiculopathy, luxation)/systemic diseases.

C. Positive if was described in what setting the patients were recruited (i.e. general practice, hospital, occupational setting).

D. Positive if the response rate was ≥75%.

E. Positive if information was presented about patient/disease characteristics of responders and non-responders or if there was no selective response.

F. Positive if a prospective design was used, also positive in case of an historical cohort in which the determinants had been measured before outcome was determined.

G. Positive if the follow-up period was at least 6 months.

H. Positive if the total number of participants was ≥75% on the last moment of follow-up compared to the number of participants at baseline.

I. Positive if demographic/clinical information (patient/disease characteristics such as age, sex and other potential prognostic predictors) was presented for completers and those lost to follow-up/drop-outs at the main moment of outcome measurement, or no selective drop-outs/lost to follow up, or no drop-outs/lost to follow-up.

J. Positive if treatment subsequent to inclusion in cohort is fully described or standardised. Also positive in case of no treatment given.

K. Positive if standardised questionnaires or objective outcome measurements of at least 1 of the following 5 outcome measures were used for each follow-up measurement: pain, general improvement, functional status, general health status or lost days of work.

L. Positive if standardised questionnaires or objective measurements were used at baseline for at least 4 of the following 8 potential prognostic factors: age, sex, pain, functional status, duration of complaints, neck complaints, physical workload, or dominant shoulder affected.

M. Positive if standardised questionnaires or objective measurements were used at baseline of at least 1 of the following 6 potential prognostic factors: depression, somatisation, distress, fear & avoidance, coping strategies, or psychosocial work-related factors (social support, psychological demands, and job decision latitude).

N. Positive if frequency, percentage or mean, median (Inter Quartile Range) and standard deviation/CI (confidence interval) were reported for the most important outcome measures.

O. Positive if frequency, percentage or mean, median (Inter Quartile Range) and standard deviation/CI were reported for the most important prognostic factors.

P. Positive if univariate crude estimates were provided for the association of a prognostic factor with outcome.

Q. Attempt is made to determine a set of prognostic factors with the highest prognostic value.

R. Positive if the number of cases in the multivariate analysis was at least ten times the number of independent variables in the analysis[40]

References

1. Luime JJ, Koes BW, Hendriksen IJM, Burdorf A., Verhagen AP, Miedema HS, Verhaar JAN. Prevalence and incidence of shoulder pain in the general population; a systematic review. In press.
2. Van der Heijden GJ. Shoulder disorders: a state-of-the-art review. Baillieres Best Pract Res Clin Rheumatol 1999;13:287-309.
3. Picavet HS, Schouten JS. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3)-study. Pain 2003;102:167-78.
4. Bongers PM. The cost of shoulder pain at work. BMJ 2001;322:64-5.
5. Van der Windt DA, Koes BW, de Jong BA, Bouter LM. Shoulder disorders in general practice: incidence, patient characteristics, and management. Ann Rheum Dis 1995;54:959-64.
6. Okkes IM, Oskam SK, Lamberts H. Van klacht naar diagnose: Episodegegevens uit de huisartspraktijk. Bussum: Coutinho, 1998.
7. Croft P, Pope D, Silman A. The clinical course of shoulder pain: prospective cohort study in primary care. Primary Care Rheumatology Society Shoulder Study Group. BMJ 1996;313:601-2.
8. Winters JC, Sobel JS, Groenier KH, Arendzen HJ, Meyboom-de Jong B. Comparison of physiotherapy, manipulation, and corticosteroid injection for treating shoulder complaints in general practice: randomised, single blind study. BMJ 1997;314:1320-5.
9. Van der Windt DA, Koes BW, Boeke AJ, Deville W, De Jong BA, Bouter LM. Shoulder disorders in general practice: prognostic indicators of outcome. Br J Gen Pract 1996;46:519-23.
10. Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. J Am Med Inform Assoc 1994;1:447-58.
11. Scholten-Peeters GG, Verhagen AP, Bekkering GE, van der Windt DA., Barnsley L, Oostendorp RA, Hendriks EJ. Prognostic factors of whiplash-associated disorders: a systematic review of prospective cohort studies. Pain 2003;104:303-22.
12. Borghouts JA, Koes BW, Bouter LM. The clinical course and prognostic factors of non-specific neck pain: a systematic review. Pain 1998;77:1-13.
13. Hudak PL, Cole DC, Frank JW. Perspectives on prognosis of soft tissue musculoskeletal disorders. Int J Rehabil Res 1998;21:29-40.
14. Altman DG. Systematic reviews of evaluations of prognostic variables. BMJ 2001;323:224-8.
15. Sackett DL, Straus SE, Richardson WS. Evidence-Based Medicine. How to Practice and Teach EBM. Guidelines. Edinburg: Churchill Livingstone, 2000.
16. Ariëns GAM, Mechelen W van, Bongers PM, Bouter LM, Wal G van der. Physical risk factors for neck pain. Scand J Work Environ Health 2000;26:7-19
17. Miranda H, Viikari-Juntura E, Martikainen R, Takala EP, Riihimaki H. A prospective study of work related factors and physical exercise as predictors of shoulder pain. Occup Environ Med 2001;58:528-34.

18. Viikari-Juntura E, Takala E, Riihimaki H, Martikainen R, Jappinen P. Predictive validity of symptoms and signs in the neck and shoulders. J Clin Epidemiol 2000;53:800-8.

19. Brox JI, Brevik JI. Prognostic factors in patients with rotator tendinosis (stage II impingement syndrome) of the shoulder. Scand J Prim Health Care 1996;14:100-5.

20. Morrison DS, Frogameni AD, Woodworth P. Non-operative treatment of subacromial impingement syndrome. J Bone Joint Surg Am 1997;79:732-7.

21. Bartolozzi A, Andreychik D, Ahmad S. Determinants of outcome in the treatment of rotator cuff disease. Clin Orthop 1994;(308):90-7.

22. Binder AI, Bulgen DY, Hazleman BL, Roberts S. Frozen shoulder: a long-term prospective study. Ann Rheum Dis 1984;43:361-4.

23. Mulcahy KA, Baxter AD, Oni OO, Finlay D. The value of shoulder distension arthrography with intraarticular injection of steroid and local anaesthetic: a follow-up study. Br J Radiol 1994;67:263-6.

24. Cassou B, Derriennic F, Monfort C, Norton J, Touranchet A. Chronic neck and shoulder pain, age, and working conditions: longitudinal results from a large random sample in France. Occup Environ Med 2002;59:537-44.

25. Macfarlane GJ, Hunt IM, Silman AJ. Predictors of chronic shoulder pain: a population based prospective study. J Rheumatol 1998;25:1612-5.

26. 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.

27. Shaffer B, Tibone JE, Kerlan RK. Frozen shoulder. A long-term follow-up. J Bone Joint Surg Am 1992;74:738-46.

28. 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.

29. Kuroda S, Sumiyoshi T, Moriishi J, Maruta K, Ishige N. The natural course of atraumatic shoulder instability. J Shoulder Elbow Surg 2001;10:100-4.

30. Chard MD, Sattelle LM, Hazleman BL. The long-term outcome of rotator cuff tendonitis; a review study. Br J Rheum 1988;27:385-9.28. 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.

31. Bjorksten MG, Talback M. A follow-up study of psychosocial factors and musculoskeletal problems among unskilled female workers with monotonous work. Eur J Public Health 2001;11:102-8.

32. Egger M, Zellweger-Zahner T, Schneider M, Kunker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. Lancet 1997;350: 326-9

33. Moher D, Pham B, Klassen TP, Schulz KF, Berlin JA, Jadad AR, Liberati A. What contributions do languages other than English make on the results of meta-analyses? J Clin Epidemiol 2000;53:964-72.

34. Moher D, Fortin P, Jadad AR, Juni P, Klassen T, Le Lorier J, Liberati A, Linde K, Penna A. Completeness of reporting of trials published in languages other than English: implications for conduct and reporting of systematic reviews. Lancet 1996;347:363-6.

35. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. Lancet 1991;337:867-72.

36. Burdett S, Stewart LA, Tierney JF. Publication bias and meta-analyses: a practical example. Int J Technol Assess Health Care 2003;19:129-34.

37. Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000;85:317-32.

38. Pincus T, Burton AK, Vogel S, Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. Spine 2002;27(5):E109-20.

39. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000;19:453-73.

40. Altman DG. Practical statistics for medical research. London: Chapman & Hall, 1991.

3

Costs of shoulder pain in primary care

Ton Kuijpers
Maurits W van Tulder
Geert JMG van der Heijden
Lex M Bouter
Daniëlle AWM van der Windt

Background Shoulder pain is common in primary care, and has an unfavourable outcome in many patients. Information on the costs associated with health care use and loss of productivity in patients with shoulder pain is very scarce.

Objective To determine shoulder pain related costs during the 6 months after first consultation in general practice.

Methods A prospective cohort study with 6 months of follow-up was conducted among 587 patients with a new episode of shoulder pain. Data on costs were collected by means of a cost diary during 6 months.

Results 84% of the patients completed all cost diaries. The mean consumption of direct health and non-health related care was low. During 6 months after first consultation for shoulder pain, the mean total costs a patient generated were € 689. Almost 50% of this concerned indirect costs, caused by sick leave from paid work. A small proportion (12%) of the population generated 74% of the total costs.

Conclusions The total costs in the 6 months after first consultation in primary care, mostly generated by a small part of the population, are not alarming. However, after 6 months 46% of the patients still reported persistent symptoms. More extensive research with a longer follow-up is needed. It is important to include patients with fractures, dislocation, or previous surgery to accurately estimate the total costs of illness for patients consulting with shoulder pain in primary care. These patients were not included in our inception cohort, but possibly generate substantial costs.

Introduction

Shoulder pain is common with a one-year prevalence ranging between 5% and 47%.[1-5] The point prevalence in the general population in The Netherlands has recently been estimated at 21%[6], while a British study reported 14%.[7] The annual incidence of shoulder pain in Dutch general practice ranges between 12 and 25/1000/year.[8;9] Shoulder pain has an unfavourable outcome in many patients. About 40 to 50% of all patients who present with a new episode of shoulder pain in primary care report persistent symptoms after 6 to 12 months.[10-12]

Musculoskeletal disorders are the second most expensive disease group for health care costs in the Netherlands, and represent 6% of the total healthcare costs.[13] Information on the costs associated with health care use and loss of productivity in patients with shoulder pain is very scarce, especially for the large majority of patients who are treated in primary health care.

We performed a cohort study among patients who presented shoulder pain to their general practitioner, and followed them for 6 months. Our objective was to determine the shoulder pain related costs during the 6 months following first consultation for their complaints in general practice.

Methods

Recruitment

Between January 2001 and June 2003, 103 general practitioners (GPs) recruited patients at first consultation for a new episode of shoulder complaints in three geographic areas in the Netherlands (Amsterdam, Groningen and Maastricht). Patients were selected if they were older than 18 years of age, and had not consulted their GP or received any form of treatment for the afflicted shoulder in the preceding three months. Sufficient knowledge of the Dutch language was required to complete written questionnaires. Exclusion criteria were severe physical or psychological conditions (i.e. fractures or dislocation in the shoulder region; rheumatic disease; neoplasm; neurological or vascular disorders; dementia).

Management of shoulder pain

The participating GPs were educated and trained to apply treatment according to the 1999 version of the Dutch guidelines for shoulder disorders issued by the Dutch College of General Practitioners.[14;15] The guidelines recommend giving information on the prognosis of shoulder pain, advice regarding provoking activities, and stepwise treatment consisting of paracetamol, Non-Steroidal Anti-Inflammatory Drugs (NSAIDs), corticosteroid injection or referral to physiotherapy. The participating GPs made the decision regarding the content of treatment based on duration and severity of pain and disability.

Table 1 Cost used in the economic evaluation

Costs	€
Direct health care costs	
General practitioner (max 10 min)	18.97
Manual therapist (max 20 min)	19.65
Other therapists (max 30 min) (physiotherapist, 'Mensendieck', 'Cesar' and occupational therapist)	19.65
Specialists (orthopaedist, neurologist, rheumatologist, and physician for rehabilitation medicine)	44.21
Hospitalisation (per day)	337.00
Direct non-health care costs	
Alternative therapists (acupuncturist, homeopath, chiropractor, and others)	As indicated by patient
Home care	8.60
Paid help at home	8.60
Help from partner/relatives/friends	8.60
Extra activities (i.e. swimming, fitness, gymnastics)	8.60
Indirect costs	
Sick leave from paid work	FCM
Sick leave from unpaid work	8.60
Unable to perform usual activities	8.60
Unable to perform hobbies	8.60

€: Euro's (€1= 1.30 US Dollars, 02/02/2005); FCM: friction cost method

Costs

Cost data were collected from a societal perspective, using a cost diary that has been shown to be valid and feasible for patient completion.[16] The diary was presented in a booklet form, containing instructions, an example of how to complete the diary, and a telephone number in case of questions. Patients were asked to complete five cost diaries during the entire follow-up period of 26 weeks, which was divided in five time periods: weeks 1-6, weeks 7-12, weeks 13-18, weeks 19-22, and weeks 23-26. Patients received a new diary by post at the beginning of a period and a return envelop for the previous diary. Patients were reminded by post after two weeks or telephoned after three weeks if they had not returned the previous diary. The cost diary included direct health care costs relevant to treatment of shoulder complaints, such as visits to a general practitioner, physiotherapist, manual therapist, occupational therapist, 'Mensendieck' or 'Cesar' exercise therapist or complementary health therapists (e.g. acupuncturist), visits to a consultant in orthopedic surgery, neurology, rheumatology, or rehabilitation medicine, and hospitalization. Direct non-health care costs included out of pocket expenses, costs of performing extra activities (i.e. swimming, fitness, gymnastics), homecare and costs for paid and unpaid help. Indirect costs included costs of loss of production due to shoulder complaints, which was measured by sick leave from paid and unpaid work, and inability to perform usual activities and hobbies. Indirect costs for paid work were calculated using the friction cost method[17;18], with a friction period of 123 days. Friction costs were based on the mean income and sex of the Dutch population.[17;18] We used a shadow price for unpaid work of € 8.60 per hour.[18;19] A complete overview of the unit costs we used is given in Table 1. Medication costs were based on the prices provided by the Royal Dutch Society for Pharmacy.[20]

Data analysis

We compared the baseline characteristics of patients who completed all cost diaries with those who did not and analysed possible differences in a multiple logistic regression analysis. Health care consumption for shoulder pain and associated costs are presented in tables, for three time periods

separately: short-term (1-6 weeks), intermediate term (7-12 weeks), and long-term follow-up (13-26 weeks). The arithmetic mean, standard deviations, maximum value (only for consumption) were computed for direct (health and non-health related), and indirect costs. Despite the skewness in the distribution of costs, it is the arithmetic mean that is the most informative measure for health care policy makers.[21;22] Measures other than the arithmetic mean provide no information about the total cost of treating all patients. We also described the costs separately for patients with persistent symptoms after 6 months and for those reporting recovery.

Results

Patients

A total of 587 patients were included in the cohort study. A total of 95 patients did not return one or more of their cost diaries. A multiple logistic regression analysis showed that patients with incomplete cost data were significantly ($p<0.10$) younger (46 years versus 52 years) and had more concomitant neck pain (38% versus 35%). The results of this study were based on the 492 patients (84%) who completed all five cost diaries. Table 2 presents their baseline characteristics.

Health care consumption and sick leave

Health care consumption and sick leave are presented in Table 3. The mean number of visits to GPs, specialists, allied health professionals, and complementary health therapists were low, as well as the mean consumption of non-health related care (home care, paid and unpaid help, extra activities). Productivity losses were not considered high as well, with a mean number of days sick leave from paid work of 2.8 (sd ± 13) over a period of 6 months. Table 3 shows that most estimations have large standard deviations.

Table 2 Baseline characteristics of patients with shoulder pain

Variable	n =492	n=61*
Demographic		
Age (years); mean (SD)	52 (14)	50 (14)
Gender: male; n (%)	245 (50)	25 (41)
Paid work; n (%)	350 (60)	38 (62)
Disease characteristics		
Duration of current shoulder complaints; n (%)		
0-6 weeks	179 (37)	18 (30)
7-12 weeks	117 (24)	15 (25)
>12 weeks	195 (40)	28 (46)
Sick leave at baseline in preceding 2 months; n (%)		
0 weeks	254 (74)	21 (57)
≥1 weeks	44 (13)	5 (13)
>1 weeks	46 (13)	11 (30)
Gradual onset; n (%)	310 (63)	36 (59)
Precipitating cause; n (%)		
Strain/overuse: usual activities; n (%)	115 (23)	22 (36)
Shoulder complaints in the past; n (%)	291 (59)	42 (69)
Neck complaints in the past; n (%)	251 (51)	41 (68)
Dominant side involved; n (%)	302 (61)	42 (69)
Comorbid psychological complaints; n (%)	42 (9)	9 (15)
Concomitant musculoskeletal complaints; n (%)		
Neck/high back	173 (35)	35 (57)
Low back pain	117 (24)	24 (39)
Upper extremity	152 (31)	21 (34)
Shoulder pain (0-10); mean (SD)	4.7 (2.3)	6.2 (1.9)
Shoulder disability (0-100); mean (SD)	60.0 (23.8)	72.0 (20.2)
Physical examination		
Pain shoulder with movement (0-18); median (IQR)	6 (4-6)	8 (6-8)
Pain neck with movement (0-18); median (IQR)	0 (0-0)	3 (0-2)
Physical factors		
Dynamic physical workload (0-5); median (IQR)	1 (0-1)	1 (1-1)
Repetitive movements; n (%)	316 (64)	43 (71)

*Subgroup of patients generating costs of >€10.000 in 6 months. SD=standard deviation; IQR=Inter quartile range

Costs

The mean costs and standard deviations are presented in Table 4. During the 6 months after first consultation for shoulder pain, the mean total costs a patient generated were € 689 (sd ± 1965). A large proportion (47%) of the total costs was due to indirect costs of productivity losses. Mean total costs due to sick leave from paid work in patients with a paid job were € 523 (sd ± 2054). Treatment by a therapist (mostly physiotherapists, Table 3) accounted for 37% of the total direct costs. One patient underwent surgery for his shoulder complaints, and generated € 2715, which reflects 1 day of hospitalization and the costs of a neck operation. The costs generated in the first six weeks (€ 276 ± 758) were almost equal to the costs generated in the final three months of the follow-up period (€ 257 ± 962). So, the highest costs per week were generated in the first 6 weeks after consultation. Table 5 shows that patients reporting persistent symptoms generated more than twice as much costs compared with patients reporting recovery after 6 months.

Figure 1 describes the distribution of the total costs in the population. A small proportion of the population (n=61; 12 %) generated more than € 1000 per patient during 6 months after first consultation. This skewness was illustrated with a median total direct costs of € 105 (Inter Quartile Range 19-317) and median total indirect costs of € 0 (IQR 0-75). These 61 patients (12%) were responsible for 74% of the total costs, which consisted for 78% of indirect costs. These small subset of the population differed considerably (>10%) at baseline from the total study population, regarding sick leave at baseline in the preceding 2 months, strain due to usual activities as precipitating cause, shoulder or neck complaints in the past, concomitant neck, high back, or back pain, intensity of shoulder pain, shoulder disability, and shoulder and neck pain at physical examination (Table 2). The prevalence of persistent symptoms after 6 months in this subgroup was 70%. Sick leave from paid work in 6 months after consultation accounted for 61% of the total costs in this subgroup.

Table 3 Consumption of healthcare resources and sick leave from work during 26 weeks after first consultation for shoulder complaints (n=492)

	week 1-6	week 7-12	week 13-26	Total (week 1-26)
Direct health care				
General practitioner [no. of visits]	1.6 ± 0.9 (6)	0.2 ± 0.6 (4)	0.3 ± 0.7 (5)	2.1± 1.5 (9)
Allied health professionals [no. of visits]				
Physiotherapist	1.5 ± 3.0 (13)	1.1 ± 2.6 (16)	1.5 ± 4.1 (28)	4.1 ± 7.6 (45)
Manual therapist	0.2 ± 1.0 (11)	0.2 ± 0.9 (9)	0.2 ± 1.1 (11)	0.5 ± 2.2 (20)
Other therapists	0.1 ± 0.5 (6)	0.0 ± 0.3 (4)	0.1 ± 0.7 (12)	0.2 ± 1.2 (19)
Specialists [no. of visits]	0.0 ± 0.3 (3)	0.0 ± 0.2 (2)	0.1 ± 0.5 (5)	0.2 ± 0.8 (7)
Hospitalisation [days]	-	-	0.0 ± 0.0 (1)	0.0 ± 0.0 (1)
Direct non-health care				
Alternative therapists [no. of visits]	0.1 ± 0.5 (6)	0.1 ± 0.5 (8)	0.2 ± 0.9 (11)	0.3 ± 1.6 (17)
Home care [hours]	0.1 ± 1.2 (22)	0.0 ± 0.8 (15)	0.2 ± 2.0 (32)	0.3 ± 3.1 (51)
Paid help [hours]	0.4 ± 3.6 (48)	0.5 ± 4.3 (60)	0.7 ± 5.7 (80)	1.7 ± 12.4 (176)
Help from partner/relatives/friends [hours]	1.1 ± 8.7 (152)	0.9 ± 8.9 (168)	1.0 ± 6.3 (72)	3.2 ± 16.1 (212)
Extra activities [hours]	0.8 ± 3.9 (44)	1.0 ± 5.3 (93)	1.4 ± 6.9 (95)	3.2 ± 11.6 (121)
Indirect				
Sick leave from paid work [days]	1.1 ± 5.0 (42)	0.6 ± 3.6 (30)	1.1 ± 6.3 (63)	2.8 ± 13.0 (123)
Sick leave from unpaid work [days]	1.9 ± 9.6 (120)	1.4 ± 9.9 (168)	1.6 ± 9.5 (104)	4.9 ± 24.2 (322)
Unable to perform usual activities [hours]	1.4 ± 7.7 (106)	0.9 ± 6.7 (108)	1.0 ± 7.3 (104)	3.2 ± 16.1 (212)
Unable to perform hobbies [hours]	2.8 ± 9.9 (120)	1.1 ± 5.7 (84)	1.3 ± 6.8 (75)	5.3 ± 16.3 (139)

Presented are group means, standard deviations and the maximum value between brackets.
Other therapists: Mensendieck, Cesar, occupational therapists, or other.
Specialists: orthopaedist; neuroligist; rheumatologist; physician for rehabilitation medicine, or other.
Alternative therapists: acupuncturist; homeopath; chiropractor; hypnotist; magnetism therapist, or other.

Table 4 Mean costs (€) made in 26 weeks after first consultation for shoulder complaints (n=492)

	week 1-6	week 7-12	week 13-26	Total week 1-26
Direct health care costs	**69 ± 66**	**32 ± 57**	**52 ± 179**	**152 ± 232**
General Practitioner	31 ± 17	4 ± 54	5 ± 14	40 ± 29
Allied health professionals	34 ± 62	25 ± 53	34 ± 87	93 ± 158
Specialists	2 ± 12	1 ± 9	5 ± 26	8 ± 36
Prescribed medication	2 ± 6	1 ± 3	1 ± 4	4 ± 10
Hospitalisation	-	-	6 ± 138	6 ± 138
Direct non-health care costs	**30 ± 102**	**26 ± 108**	**41 ± 146**	**98 ± 293**
Alternative therapists	5 ± 33	3 ± 28	8 ± 55	16 ± 94
Over-the-counter medication	2 ± 9	1 ± 9	4 ± 51	7 ± 58
Out-of-pocket expenses	2 ± 16	1 ± 8	1 ± 6	4 ± 21
Home care	1 ± 10	0.4 ± 7	2 ± 17	3 ± 27
Paid help	4 ± 31	5 ± 37	6 ± 49	15 ± 107
Help from partner/relatives/friends	10 ± 75	7 ± 76	8 ± 54	25 ± 168
Extra activities	7 ± 34	8 ± 45	12 ± 59	28 ± 100
Total direct costs	**99 ± 130**	**58 ± 134**	**93 ± 244**	**250 ± 408**
Sick leave from paid work	125 ± 654	69 ± 461	130 ± 790	324 ± 1635
Sick leave from unpaid work	16 ± 83	12 ± 85	13 ± 82	42 ± 208
Unable to perform usual activities	12 ± 66	7 ± 57	9 ± 63	28 ± 139
Unable to perform hobbies	24 ± 85	10 ± 49	12 ± 59	46 ± 140
Total indirect costs	**177 ± 711**	**99 ± 505**	**164 ± 829**	**439 ± 1751**
Total costs	**276 ± 758**	**156 ± 564**	**257 ± 962**	**689 ± 1965**

Presented are group means, standard deviations

Table 5 Mean costs (€) and standard deviations in patients with persistent symptoms versus those in recovered patients after 6 and 26 weeks (n=492)

	week 1-6		week 1-26	
	Persistent symptoms	*Recovered*	*Persistent symptoms*	*Recovered*
Direct health care costs	82 ± 71	48 ± 49	206 ± 301	107 ± 133
Direct non-health care costs	37 ± 115	20 ± 82	133 ± 316	70 ± 274
Total direct costs	120 ± 144	68 ± 98	339 ± 477	177 ± 325
Indirect costs	208 ± 727	127 ± 679	628 ± 2049	279 ± 1458
Total costs	**328 ± 780**	**195 ± 696**	**966 ± 2334**	**457 ± 1573**

Presented are group means, and standard deviations.

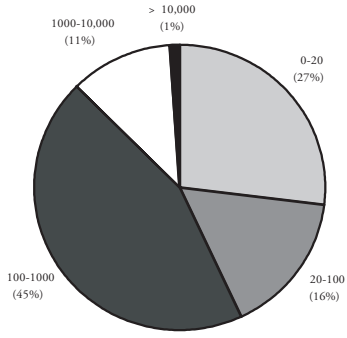


Figure 1 Distribution of the total costs (€) in the population (n=492)

Discussion

The present study is the first to evaluate the overall costs generated by patients presenting with shoulder pain in primary care. Healthcare consumption and sick leave did not seem to be alarmingly high in this primary care population with mean total costs of € 689 per patient during the 6 months after first consultation. A small part (12%) of the population accounted for 74% of the total costs.

The response rate was high (84%) and differences at baseline between completers (all 5 five cost diaries) and non-completers (<5 cost diaries) were small. We expect that the differences in age and concomitant neck pain between completers and non-completers will not have substantially influenced our cost estimates.

Although there was a high percentage of patients who reported persistent symptoms after 6 weeks (70%) and 6 months (46%) in this population[23], there was little health care consumption and shoulder pain related sick leave in this study. As a consequence shoulder pain related costs per patient were low. Only one patient reported 130 days of sick leave in the 6 months after consultation, which was more than the friction period of 123 days.

An explanation for the modest health care costs could be that general practitioners stick to the interventions recommended in the Dutch guidelines for shoulder disorders[14;15] (wait-and-see policy with pain medication, followed by injections), which are relatively inexpensive. The costs of physiotherapy represented a relatively large proportion of the direct health care costs, but accounted for only 14% of the total costs, as few patients were referred for therapy (Table 4). Indirect costs accounted for a large proportion (47%) of the total costs. Nevertheless, the total number of days sick leave per patients was small (2.8 days) over a period of 6 months. Possibly, factors such as shoulder pain, sleeping problems, or loss of function have caused loss of productivity in patients without sick leave from paid work. Our study does not provide information on the actual loss of productivity among those who kept on working regardless of the shoulder pain.

Similar to studies on low back pain[24], in this study a small proportion of the population (n=61; 12%) caused a substantial part (74%) of the total costs. In this subgroup sick leave from paid work accounted for 61% of the total costs. In a prognostic study we found the following factors predicting sick leave from paid work in 6 months after consultation: sick leave in the preceding 2 months, shoulder pain, precipitating cause: usual activities and concomitant psychological complaints.[25] Table 2 shows that the subgroup who generated most costs, not surprisingly, also showed higher scores on these variables compared to the total population (Table 2). Table 2 also shows substantial differences between the groups regarding shoulder or neck pain at physical examination, and concomitant back pain at baseline. This seems plausible as these factors were shown to be of predictive value for persistent shoulder pain after 6 weeks and 6 months in our cohort.[23]

In this study we were able to include a follow-up of 26 weeks. It is possible that prolonged and recurrent pain episodes generate additional costs for more expensive care, e.g. diagnostic imaging and surgical treatment, including hospitalization. Given the poor prognosis of shoulder pain (approximately 40% of patients report persistent symptoms after 12 to 18 months[10-12]) higher health care costs and productivity losses may be expected when follow-up times are longer. In the 6 months following first

consultation, few costs were made due to referrals to other health care providers, additional diagnostic procedures, or surgery. We expect these kind of health care expenses to occur in the long-term in a small subset of the population. In our inception cohort patients with fractures, dislocation, or previous surgery were not included. These patients are not included in this study, but may generate substantial costs in current practice.

Information on the costs associated with health care use and loss of productivity in patients with shoulder pain is very scarce, and therefore a comparison with other studies is difficult. In the framework of this cohort study a randomised controlled trial on the effectiveness of manipulative therapy has been carried out.[26] This trial was similar regarding outcome assessments and length of follow-up. The control group of this trial (n=71), who also received usual care according to the Dutch general practice guidelines, generated slightly lower costs (mean total costs € 555) compared to our cohort.

In conclusion, the total costs in the 26 weeks after first consultation for shoulder pain, mostly generated by a small part of the population, are not alarming. However, after 26 weeks 46% of the patients still reported persistent symptoms. More extensive research with a longer follow-up to monitor these patients, is needed. These patients possibly generate substantial costs. Given the high incidence of shoulder pain (12/1000/year)[8;9] in general practice the total costs to society could be substantial. It may be important to include patients with fractures, dislocation, or previous surgery to estimate the total costs of illness for patients consulting with shoulder pain in primary care.

Reference List

1. Luime JJ, Koes BW, Hendriksen IJ, et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. Scand J Rheumatol 2004;33:73-81.
2. Van der Heijden GJMG. Shoulder disorders: a state-of-the-art review. Baillieres Best Pract Res Clin Rheumatol 1999;13:287-309.
3. Bernard B, Sauter S, Fine L, Petersen M, and Hales T. Job Task and Psychosocial Risk Factors for Work-Related Musculoskeletal Disorders Among Newspaper Employees. Scand J Work Environ Health 1994;20:417-26.
4. Skov T, Borg V, and Orhede, E. Psychosocial and Physical Risk Factors for Musculoskeletal Disorders of the Neck, Shoulders, and Lower Back in Salespeople. Occup Environ Med 1996;53:351-6.
5. Trinkoff AM, Lipscomb JA, Geiger-Brown J, and Brady B. Musculoskeletal Problems of the Neck, Shoulder, and Back and Functional Consequences in Nurses. Am J Ind Med 2002;41:170-8.
6. Picavet HS, Schouten JS. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3)-study. Pain 2003;102:167-78.
7. Bongers PM. The cost of shoulder pain at work. BMJ 2001;322:64-65.
8. Okkes IM, Oskam SK, Lamberts H. Van klacht naar diagnose: Episodegegevens uit de huisartspraktijk. Bussum. Coutinho, 1998.
9. Van der Windt DAWM, Koes BWK, Jong BAD, et al. Shoulder disorders in General Practice: incidence, patients characteristics, and management. Annals of Rheumatic Diseases 1995;54:959-64.
10. Croft P, Pope D, Silman A. The clinical course of shoulder pain: prospective cohort study in primary care. BMJ 1996;313:601-2.
11. Van der Windt DAWM, Koes BWK, Boeke AJP, et al. Shoulder disorders in general practice: prognostic indicators of outcome. Br J Gen Pract 1996;46:519-23.
12. Winters JC, Sobel JS, Groenier KH, et al. The long-term course of shoulder complaints: a prospective study in general practice. Rheumatology 1999;38:160-3.
13. Meerding WJ, Bonneux L, Polder JJ, Koopmanschap MA, van der Maas PJ. Demographic and epidemiological determinants of healthcare costs in Netherlands: cost of illness study. BMJ 1998;317:111-15.
14. The Dutch College of General Practitioners (NHG) Practice Guideline. Accessed at http://nhg.artsennet.nl/upload/104/guidelines2/E08.htm on 1 september 2003.
15. Winters JC, Jongh AC, van der Windt DAWM, et al. Practice Guideline "Shoulder complaints' [In Dutch: NHG-Standaard Schouderklachten]. Huisarts en Wetenschap. 1999;42:222-31.
16. Goossens ME, Rutten-van Molken MP, Vlaeyen JW, van der Linden SM. The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. J Clin Epidemiol 2000;53:688-95.
17. Oostenbrink JB, Koopmanschap MA, Rutten FFH. Manual costs research [In Dutch: Handleiding voor kostenonderzoek, methoden en richtlijnprijzen voor economische evaluaties in de gezondheidszorg]. Amstelveen: College voor zorgverzekeringen, 2000.

18. Oostenbrink JB, Koopmanschap MA, Rutten FF. Standardisation of costs: the Dutch Manual for Costing in economic evaluations. Pharmacoeconomics 2002;20:443-54.
19. Taxe report: Z-Index [In Dutch]. 2000. The Hague.
20. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Clin Oral Investig 2003;7:2-7.
21. Barber JA, Thompson SG. Analysis and interpretation of cost data in randomised controlled trials: review of published studies. BMJ 1998;317:1195-00.
22. Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? BMJ 2000;320:1197-00.
23. Kuijpers T, Van der Windt DAWM, Boeke AJ, Twisk J, Vergouwe Y, LM Bouter, Van der Heijden GJMG. Clinical prediction rules for the prognosis of shoulder pain in general practice. Pain, in press.
24. Cats-Baril WL, Frymeyer JW. The economics of spinal disorders. In: Frymeyer JW, editor. The adult spine: principles and practice. New York: Raven Press, 1991.
25. Kuijpers T, Van der Windt DAWM, Van der Heijden GJMG, Twisk J, Vergouwe Y, Bouter LM. A prediction rule for shoulder pain related sick leave. Submitted.
26. Bergman GJD, Winters JC, Groenier KH, Pool JJ, Meyboom-de Jong B, Postema K et al. Manipulative therapy in addition to usual medical care for patients with shoulder dysfunction and pain: a randomized, controlled trial. Ann Intern Med 2004;141:432-39.

4

Clinical prediction rules for the prognosis of shoulder pain in general practice

Ton Kuijpers
Daniëlle AWM van der Windt
A Joan P Boeke
Jos WR Twisk
Yvonne Vergouwe
Lex M Bouter
Geert JMG van der Heijden

Pain, in press

ABSTRACT

Background Shoulder pain is common in primary care, and has an unfavourable outcome in many patients. Information about predictors of outcome is scarce and inconsistent. The objective of this study was to develop clinical prediction rules for calculating the absolute risk of persistent shoulder symptoms for individual patients, 6 weeks and 6 months after the first consultation in general practice.

Methods A prospective cohort study with 6 months follow-up was carried out in three geographic areas in The Netherlands. In this study 587 patients with a new episode of shoulder pain were included. The main outcome measure was persistent symptoms at 6 weeks and 6 months, perceived by the patient. Potential predictors included the results of a physical examination, sociodemographic variables, disease characteristics (duration of symptoms, pain intensity, disability, comorbidity), physical activity, physical workload, and psychosocial factors.

Results Response rates to the follow-up questionnaires were 83% at 6 weeks and 92% at 6 months. A longer duration of symptoms, gradual onset of pain, and high pain severity at presentation were consistently associated with persistent symptoms at 6 weeks and 6 months. The discriminative validity of our prediction rules was satisfactory with area under the curves of 0.74 (95% CI 0.70; 0.79) at 6 weeks and 0.67 (95% CI 0.63; 0.71) at 6 months. The performance of our rules needs to be tested in other populations of patients with shoulder pain to enable valid and reliable use of the rules in everyday clinical practice.

Introduction

Shoulder pain is common with a one year prevalence ranging between 5% and 47%.[1-5] The point prevalence in the general population in The Netherlands has recently been estimated at 17%.[6] The annual incidence of consultation for a new episode of shoulder pain in Dutch general practice ranges between 12 and 25/1000/year.[1;6-8]

Shoulder pain has an unfavourable outcome in many patients. Only about 50% of all new episodes of shoulder pain presented in primary care show complete recovery within six months[9-11], after one year this proportion increases to only 60%.[10] Knowing more about the prognostic value of clinical, psychosocial, and occupational factors in patients with shoulder disorders will help to provide patients with adequate information regarding the most likely course of their symptoms. Such information may also support decisions regarding treatment and referral of patients.

In a systematic review of the literature we summarized the available evidence from 16 studies regarding predictors of outcome of shoulder pain.[12] Only six studies were of relatively high quality. In a primary care population strong evidence for predicting poor outcome was only found for 'high pain intensity'. For any other variable, including psychosocial variables, convincing evidence for their predictive value is lacking. We performed a cohort study among patients with shoulder pain consulting their general practitioners, and followed them for 6 months. The objective of this study was to determine which combination of factors predicts the outcome of an episode of shoulder pain 6 weeks and 6 months after the first consultation in a general practice population. Our aim was to develop a clinical prediction rule for calculating the absolute risk of persistent symptoms for individual patients in general practice.

Methods

Recruitment

Between January 2001 and June 2003, 103 general practitioners (GP) recruited patients at first consultation for a new episode of shoulder pain in three geographic areas in the Netherlands (Amsterdam, Groningen and

Maastricht). The primary reason for consultation had to be shoulder pain. In this study shoulder pain was defined according to the 1999 version of the Dutch guidelines for shoulder complaints, issued by the Dutch College of General Practitioners.[13;14] In the guideline shoulder pain is characterised as pain in the deltoid and upper arm region. GPs used this definition to select patients with shoulder pain for our study.

Patients were selected by their GP if they were 18 years or older of age, and had not consulted their GP or received any form of treatment for the afflicted shoulder in the preceding 3 months. GPs were instructed to select consecutive patients. Sufficient knowledge of the Dutch language was required to complete written questionnaires. Exclusion criteria were severe physical or psychological conditions (i.e. fractures or luxation in the shoulder region; rheumatic disease; neoplasm; neurological or vascular disorders; dementia). Data collection was approved by the Medical Ethics Committee of the VU University Medical Center, Amsterdam, The Netherlands.

Management of shoulder pain

All patients received standardised treatment according to the 1999 version of the Dutch guidelines for shoulder complaints issued by the Dutch College of General Practitioners.[13;14] The guidelines recommend giving information on the prognosis of shoulder pain, advice regarding provoking activities, and stepwise treatment consisting of paracetamol, Non-Steroidal Anti-Inflammatory Drugs (NSAIDs), corticosteroid injection or referral for physiotherapy. The GP made the decision regarding the content of treatment based on duration and severity of pain and disability. The participating general practitioners were educated and trained to apply treatment according to this guideline.

Prognostic factors

Within 10 days after they had consulted the GP participants gave written informed consent and completed an extensive baseline questionnaire. The questionnaire contained questions on socio-demographic variables, disease

characteristics (i.e. pain intensity, disability, duration of complaints, onset, and comorbidity), physical activity, physical workload and psychosocial factors, and had a physical examination by a trained assistant at baseline.

The sociodemographic variables and disease characteristics were measured using checklist which mainly consisted of yes or no questions. Physical activity was measured with a single question (less/equally/more active than others). We measured physical workload with a self-constructed scale of 5 questions (yes/no) concerning pushing and pulling, lifting weights, working with hands above shoulder level, and the use of vibrating tools on at least two days a week (total score 0-5, Crohnbach's α= 0.74). Repetitive movements, on at least two days a week, was also measured with a single question answered with yes or no.

The psychosocial factors coping, anxiety, depression, somatization, distress, fear-avoidance beliefs, and kinesiophobia were measured with widely used standardized questionnaires. Coping was assessed with the 43-item Pain Coping and Cognition List (PCCL)[15], consisting of the subdomains catastrophizing (1-6 points, Crohnbach's α= 0.63), coping with pain (1-6 points, Crohnbach's α= 0.83), internal (1-6 points, Crohnbach's α= 0.76) and external locus of control (1-6 points, Crohnbach's α= 0.65). Anxiety (0-24 points, Crohnbach's α= 0.77), depression (0-12 points, Crohnbach's α= 0.90), somatization (0-32 points, Crohnbach's α= 0.82), and distress (0-32 points, Crohnbach's α= 0.92), were measured with the 50-item Four-Dimensional Symptom Questionnaire (4DSQ).[16;17] Fear-avoidance beliefs were assessed using the 4-item physical activity subscale of the Fear Avoidance Beliefs Questionnaire (FABQ; 0-24, Crohnbach's α= 0.73).[17;18] Kinesiophobia, finally, was measured using two items (no. 1 and no 9.) of the Tampa Scale for Kinesiophobia (TSK; 0-12 points, Crohnbach's α= 0.82).[20;21] The questionnaire also included a general one-item question regarding the presence (yes/no) of any psychological problems (e.g. distress, depression, anxiety).

Function of the shoulder joint and cervicothoracic spine were tested during a physical examination. For the glenohumeral joint active and passive abduction, passive exorotation[22], and shoulder impingement[23] were tested. Two alternative functional tests, HIB (Hand-in-back) and HIN

(Hand-in-neck)[24;25] measured on a 7-point scale (score 0 = very poor range of motion, score 7 = full range of motion) were performed as well. The assistant made an estimation of the range of motion in degrees (°).

During all mobility tests self-reported pain was assessed on a 4-point scale (0 = no pain; 3 = severe pain). A factor analysis on the results of a physical examination in a similar population of patients with shoulder pain resulted in four factors: shoulder mobility, shoulder pain, neck mobility, and neck pain.[22]

The factor 'shoulder mobility' consisted of 6 mobility tests: HIB, HIN, active abduction, passive abduction, exorotation, and impingement. For calculation of the sum score (0-18 points) variables were recoded into a 4-point scale, with 0 reflecting full range of motion and 3 points reflecting very poor range of motion. HIB/HIN scores were recoded as: score 7 = 0; score 5 and 6 = 1; score 3 and 4 = 2; score 1 and 2 = 3. Abduction (active and passive) was recoded as 170-180° = 0; 140-170° = 1; 90-140° = 2; 0-90° =3. Exorotation was recoded as >80° = 0; 70-80° = 1; 50-70° = 2; <50° =3. During the impingement test pain was measured (0 = no pain; 3 = severe pain). The factor 'shoulder pain' (0-18 points) consisted of the sum of the pain scores during the mobility tests.

The factor 'neck mobility' (0-4 points) consisted of rotation of the cervicothoracic spine in neutral, flexed, and extended position and lateral bending. These range of motion tests were scored as (1= decreased range of motion, and 0 = no decreased range of motion). The factor 'neck pain' (0-18 points) consisted of the sum of the pain scores during flexion and extension of the neck, rotation in a neutral, flexed and extended position, and lateral bending.

Outcome measurements

The outcome was measured by postal questionnaires at 6 weeks, 3 and 6 months. We restricted the length of the follow up period to 6 months because the recovery rate, which is about 50% after 6 months[9-11], only slightly increases thereafter. This means that little information can be gained after 6 months follow-up. Our primary outcome measure 'Patient

perceived recovery' was measured on an 8-point scale. Patients who did not report full recovery or very much improvement were denoted as having "persistent symptoms".[26;27] Patients who did not reply at 6 weeks were re-contacted at 3 and 6 months. Secondary outcome measures were shoulder disability, measured with the 16-item shoulder disability questionnaire (SDQ; 0-100)[28], pain (0-10 numeric rating scale)[26], and severity of the main complaint (0-10 numeric rating scale).[29] We studied the relationship between our primary and secondary outcome measures to determine if patients with persistent symptoms after 6 weeks and 6 months showed higher levels of pain and disability.

Analysis

Missing values of patient characteristics were imputed (approx. 1% of all required values at both 6 weeks and 6 months). Imputation was based on the correlation between each variable with missing values with the other patient characteristics. Univariable logistic regression analyses were performed for all potential prognostic factors with our primary outcome measure, i.e. persistent symptoms, at either short term (6 weeks) or long term (6 months). The linearity of the associations of continuous variables with outcome was studied. Factors were categorized if they did not show a linear association with the outcome. Variables that had a statistically significant association with the outcome (p-value ≤0.20) were selected as candidate predictors for the multivariable analysis. Nor more than one independent variable per ten events was included in the multivariable analysis.[30;31] We presented the univariable ORs along with the 95% confidence intervals, as well as with p-values to enable the reader to choose alternative statistical significance levels for the selection of variables for the multivariable analysis.

Separate prediction models were developed for persistent symptoms at short term and long term. A second selection step was performed in the multivariable model, that contained all candidate predictors with stepwise backward selection. Variables with the lowest predictive value were deleted from the model until further elimination of a variable resulted in a statistically significant lower model fit estimated with the log likelihood

ratio test ($p<0.20$).

Bootstrapping techniques were used to study the internal validity of the final prediction model, i.e. to adjust the estimated regression coefficients for overfitting and the model performance for overoptimism.[32;30] The model's performance obtained after bootstrapping can be considered as the performance that can be expected in similar future patients. Random bootstrap samples were drawn with replacement (100 replications) from the full data set. The multivariable selection of variables was repeated within each bootstrap sample. All analyses were performed using S-plus 6.1 (Insightful Corp., Seattle, WA, USA).

Evaluation of the model

The reliability of the multivariable model was determined by use of the Hosmer-Lemeshow goodness-of-fit statistic.[33] Calibration of the model predictions was assessed by plotting the predicted individual probabilities against the observed individual probabilities for persistent symptoms. For this, patients were grouped into deciles according to their predicted probability for persistent symptoms according to the model. The prevalence of the endpoint within each decile represents the observed individual probability. The area under the receiver-operating characteristic curve (ROC) was used to assess the discriminative ability of the model. The ROC-curve plots the true positive rate (sensitivity) against the false-positive rate (1-specificity) at any given cut-off value. The curve illustrates the ability of the model to discriminate between patients with and without persistent symptoms at subsequent cut-off points along the range of the predicted probabilities. An area under the curve (AUC) of 0.5 indicates no discrimination above chance, whereas an AUC of 1.0 indicates perfect discrimination.

From a prediction model to an individual patient risk

We developed a clinical prediction rule[34-36] for outcome at 6 weeks and 6 months, to provide an estimate for individual patients of their absolute risk of persistent symptoms. The probability (P) of persistent symptoms

was predicted by $P=1/[1+ \exp - (a_0 + b_1x_1 + \ldots\ldots + b_jx_j)]$. The status of a patient for any dummy or binary variable included in the prediction rule can be either 0 or 1, while for a (semi) continuous variable it takes the actual observed value.

Score charts

To facilitate the calculation of an individual patient's risk, we developed score charts. We divided the regression coefficients by the lowest coefficient and rounded them to the nearest integer to form the scores for the predictors. The sum of the scores correspond to a risk of poor outcome.

Results

Study population and follow-up

At baseline 587 patients were questioned and physically examined. Table 1 lists the baseline characteristics of the participants. At 6 weeks 487 (83%) and at 6 months 538 (92%) patients returned the postal questionnaire. The drop-outs at 6 weeks and 6 months were younger than the responders (mean difference 4 years and 6 years, respectively). The drop-outs at 6 months showed more often an acute onset (49% versus 36%), and less repetitive movements in their work (26% versus 36%) at baseline in comparison with the responders.

At 6 weeks 70% (n=340) and at 6 months 46% (n=249) patients reported persistent symptoms. Of these 249 patients, only 22 reported that symptoms had recurred after initial recovery at 6 weeks. Table 2 shows that patients with persistent symptoms reported also more pain, more shoulder disability and higher severity of the main complaint.

Management of shoulder pain

At baseline most patients (n=423, 72%) received a wait and see policy, paracetamol, or NSAIDs. Furthermore, 68 patients (12%) received an injection with corticosteroid, 58 patients (10%) were referred for physiotherapy and 28 patients (6%) received other therapies.

Table 1 Baseline characteristics of patients with shoulder disorders (n=587) and univariable associations with persistent symptoms at 6 weeks and 6 months

Variable	n(%)	6 weeks OR	95% CI	p[a]	6 months OR	95% CI	p[a]
Demographic							
Age (years); mean (SD)	51 (14)	1.0	1.0, 1.0	0.32	1.0[#]	1.0, 1.0	0.29
Gender: male	292 (50)	1.3	0.9, 2.0	0.14	0.9	0.6, 1.3	0.65
Education				0.04			0.12
Low*	210 (36)						
Middle	234 (40)	1.0	0.6, 1.6		0.7	0.5, 1.0	
High	135 (23)	0.6	0.3, 0.9		0.7	0.4, 1.0	
Disease characteristics							
Duration of complaints				<0.001			<0.001
0-6 weeks*	205 (35)						
7-12 weeks	139 (24)	2.3	1.4, 3.9		1.8	1.2, 2.9	
>3 months	242 (41)	5.4	3.3, 8.9	<0.001	3.5	2.3, 5.2	<0.001
Gradual onset (vs. acute)	363 (62)	2.9	1.9, 4.3		2.2	1.5, 3.0	
Precipitating cause				0.59			0.43
Unexpected movement	33 (6)	0.8	0.4, 1.8	0.21	1.3	0.6, 2.8	0.04
Strain/overuse: unusual activities	96 (16)	3.6	1.1, 12.2	0.02	0.6	0.4, 1.0	0.25
Strain/overuse: usual activities	138 (24)	1.8	1.1, 3.1	0.03	1.3	0.8, 1.9	0.03
Injury	33 (6)	3.6	1.1, 12.2	0.67	2.4	1.1, 5.5	0.82
Sport injury	29 (5)	1.2	0.5, 3.2	0.03	0.9	0.4, 2.0	0.67
Unknown	239 (41)	0.7	0.4, 1.0	0.16	0.9	0.7, 1.3	0.10
Shoulder complaints in the past	348 (62)	1.3	0.9, 2.0	0.00	1.3	0.9, 1.9	0.04
Neck complaints in the past	296 (51)	1.9	1.3, 2.8	0.24	1.4	1.0, 2.0	0.29
Dominant side involved	362 (62)	1.3	0.9, 1.9	0.01	1.2	0.9, 1.7	0.19
Comorbid psychological complaints	55 (9)	3.3	1.3, 8.7		1.5	0.8, 2.7	

Table 1 continued

Variable	n(%)	6 weeks OR	95% CI	p[a]	6 months OR	95% CI	p[a]
Concomitant musculoskeletal complaints							
Neck/high back	209 (36)	1.7	1.1, 2.6	0.01	1.6	1.1, 2.2	0.01
Low back pain	139 (24)	1.5	0.9, 2.3	0.13	2.2	1.5, 3.3	<0.001
Upper extremity	174 (30)	2.0	1.2, 3.8	<0.001	1.7	1.2, 2.4	0.01
Lower extremity	177 (30)	1.2	0.8, 1.9	0.44	1.7	1.2, 2.5	<0.001
Shoulder pain (0-10); mean (SD)	4.8 (2.3)	1.3	1.1, 1.4	<0.001	1.2	1.1, 1.3	<0.001
Shoulder disability (0-100); mean (SD)	59.9 (24.2)	1.0	1.0, 1.0	<0.001	1.0	1.0, 1.0	<0.001
Physical examination							
ROM shoulder (0-18); mean (SD)	6.8 (4.3)	1.1	1.0, 1.2	0.01	1.7[#]	1.0, 2.2	0.22
Pain shoulder with movement (0-18); median (IQR)	4 (2-4)	1.1	1.1, 1.7	<0.001	1.1	1.1, 1.2	<0.001
ROM neck (0-4); median (IQR)	0 (0-0)	1.1	1.0, 1.3	0.10	0.9[#]	0.6, 1.4	0.53
Pain neck with movement (0-18); median (IQR)	0 (0-0)	1.2	1.1, 1.3	<0.001	1.1	1.0, 1.2	0.01
Physical factors							
Dynamic physical workload (0-5); median (IQR)	1 (1-2)	1.2	1.0, 1.4	0.02	1.0[#]	0.6, 1.5	0.11
Repetitive movements	384 (65)	2.1	1.4, 3.1	<0.001	1.2	0.8, 1.7	0.33
Physical activity in comparison to others				0.03			0.20
more active*	126 (39)						
equally active	245 (42)	0.6	0.4, 0.9		0.9	0.6, 1.3	
less active	110 (19)	1.2	0.7, 2.2		1.4	0.9, 2.3	

Table 1 continued

Variable	n(%)	6 weeks OR	95% CI	p[a]	6 months OR	95% CI	p[a]
Psychosocial factors							
Coping (mean, SD)							
Catastrophizing (1-6)	2.2 (0.8)	1.4	1.1, 1.8	0.02	1.4[#]	0.7, 2.6	0.42
Coping with pain (1-6)	3.1 (1.0)	1.0[#]	0.2, 4.2	0.96	2.2[#]	0.6, 9.0	0.21
Internal locus of control (1-6)	3.3 (0.9)	0.8[#]	0.2, 3.9	0.45	1.4[#]	0.4, 5.3	0.62
External locus of control (1-6)	3.2 (0.9)	0.6[#]	0.1, 3.1	0.73	1.3[#]	0.7, 1.3	0.32
4DSQ (median, IQR)							
Distress (0-32)	0 (0-0)	2.2[$]	0.7, 6.6	0.15	2.6[$]	1.2, 5.8	0.02
Depression (0-12)	0 (0-0)	3.0[#]	0.4, 25.2	0.52	3.0[#]	0.6, 15.4	0.36
Anxiety (0-24)	0 (0-0)	1.8[$]	0.4, 8.0	0.46	1.2[#]	0.2, 5.9	0.85
Somatization (0-32)	0 (0-2)	5.2[$]	1.2, 22.4	0.03	2.5[$]	1.1, 5.4	0.02
Fear-avoidance (0-24); mean (SD)	14.1 (5.6)	1.0	1.0, 1.1	0.74	1.0	1.0, 1.0	0.71
Kinesiophobia (0-12); median (IQR)	2 (0-2)	0.9[#]	0.4, 2.0	0.99	1.5[#]	0.7, 3.0	0.26

SD= standard deviation; IQR = Inter quartile range; ROM = Range of Motion; 4DSQ = Four-dimensional symptom questionnaire. [a]Variables with a univariable p-value ≤0.20 were selected for the multivariable analysis of persistent symptoms at 6 weeks and 6 months. *Reference category. [#]In case of non-linear associations continuous variables were divided into categories. The table presents the Odds Ratio (OR) for the highest versus lowest category. [$]Variable was dichotomised.

Prognostic factors

Table 1 also presents the univariable association of potential predictors with outcome at 6-weeks and 6-months follow-up. Given the fact that median baseline scores on distress, anxiety and somatization were very low, scores on these psychological factors were dichotomized. Variables which showed an univariable association ($p \leq 0.20$) were selected for the backward stepwise selection analysis. Table 3 presents the variables included in the prediction models for persistent symptoms at 6 weeks and 6 months after backward stepwise selection ($p \leq 0.20$). A longer duration of symptoms at baseline, gradual onset of shoulder complaints, and higher pain intensity were associated with a poorer prognosis at both 6 weeks and 6 months. Furthermore, concomitant psychological complaints, repetitive movements, and increasing neck pain scores at physical examination were associated with persistent symptoms at 6 weeks. A poor prognosis at 6 months was additionally predicted by concomitant back pain and increasing shoulder pain scores at physical examination.

Table 2 Secondary outcome measures (mean; SD) for patients with and without persistent symptoms at 6 weeks and 6 months

	6 weeks		6 months	
Persistent symptoms	Yes	No	Yes	No
Pain (0-10)	4.3 (2.1)	0.5 (0.9)	4.1(2.3)	0.4 (1.1)
Shoulder disability (SDQ) (0-100)	53.0 (25.5)	10.4 (16.6)	52.2 (26.7)	5.9 (14.5)
Severity of main complaint (0-10)	4.8 (2.6)	0.8 (1.5)	5.0 (2.8)	0.6 (1.3)

Evaluation of the models

The reliability of the models was adequate, according to the Hosmer-Lemeshow statistic, with a p-value of 0.51 for the model at 6 weeks and 0.16 at 6 months. Figure 1 shows the calibration of the predictions. The predicted and observed probabilities are rather close to the 45° line, demonstrating good calibration of the predictions by the two models. The AUCs for the models at 6 weeks and 6 months were 0.74 (95% CI 0.70; 0.79) and 0.67 (95% CI 0.63; 0.71). The predicted risks of persistent symptoms are widely distributed (Figure 2).

Score charts

Figure 3 shows the score charts for calculating the risk of persistent symptoms at the short and long term. For instance, a patient with shoulder complaints for 3 weeks at baseline with a gradual onset of symptoms, and a shoulder pain score of 1 point, has a prognostic score of 8 points for the short term and 12 points in the long term, which implies 40 to 50% risk of persistent symptoms at 6 weeks and 20 to 30% at 6 months.

Table 3 Multivariable model with predictors of persistent shoulder symptoms (yes/no) at 6 weeks and 6 months after stepwise backward selection

Predictor	Scale	OR	95% CI
6 weeks (n=486)			
Duration of complaints			
0-6 weeks*			
7-12 weeks	(yes/no)	1.9	1.1-3.3
>3 months	(yes/no)	2.6	1.5-4.4
Gradual onset	(yes/no)	1.8	1.1-2.9
Concomitant psychological complaints	(yes/no)	2.3	0.9-6.4
Repetitive movements	(yes/no)	2.0	1.2-3.1
Shoulder pain	(0-10)	1.1	1.0-1.2
Neck pain score at physical examination	(0-18)	1.1	1.0-2.7
6 months (n=538)			
Duration of complaints			
0-6 weeks*			
7-12 weeks	(yes/no)	1.4	0.9-2.3
>3 months	(yes/no)	1.9	1.2-3.0
Gradual onset	(yes/no)	1.4	1.0-1.8
Concomitant low back pain	(yes/no)	1.6	1.1-2.5
Shoulder pain	(0-10)	1.1	1.0-1.2
Shoulder pain score at physical examination	(0-18)	1.0	1.0-1.1

*Reference category

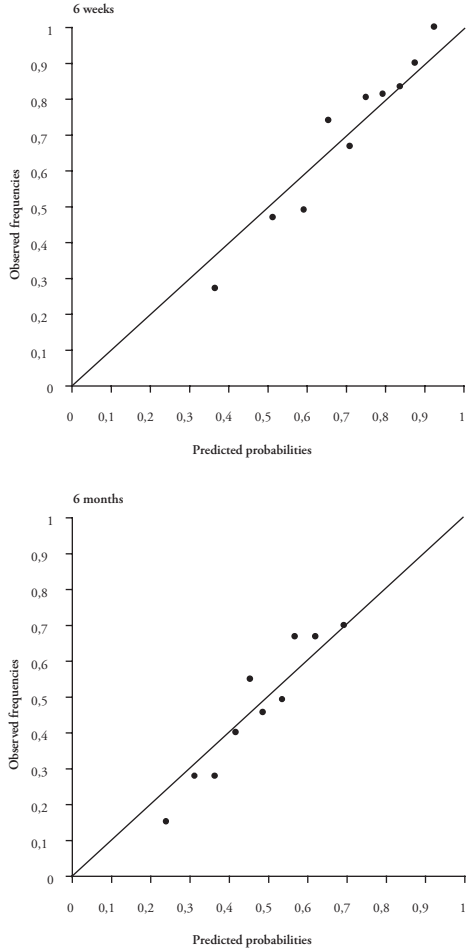


Figure 1 Calibration plots showing the observed frequencies versus the predicted probabilities for persistent symptoms at 6 weeks and 6 months

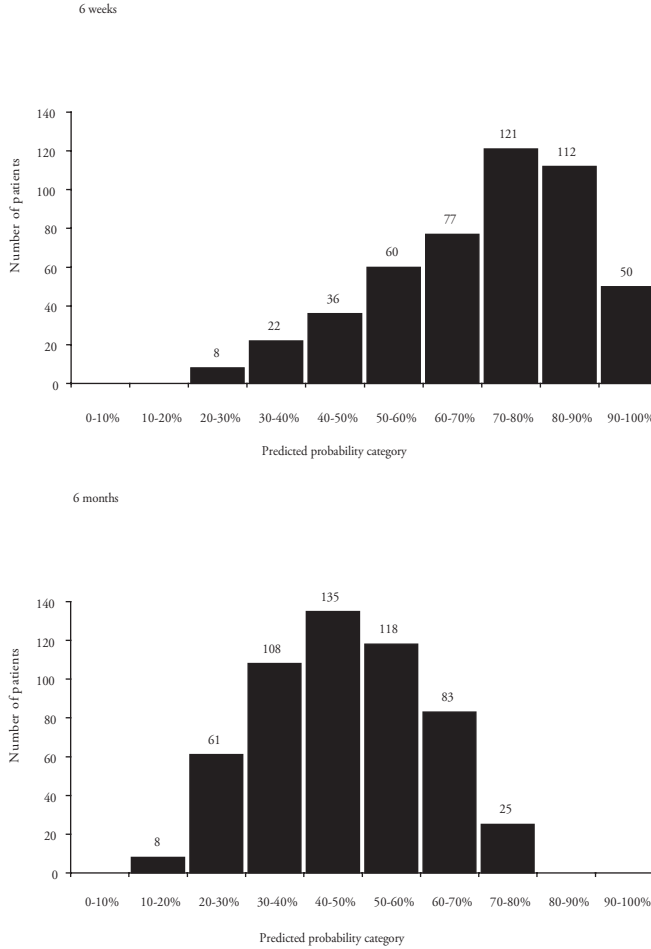


Figure 2 Number of patients in risk categories for persistent symptoms of the score charts for 6 weeks (n=486) and 6 months (n=538)

Instruction
If a predictor is scored positively, the given weight needs to be filled in. Subsequently the scores are added to calculate the 'Total score'. From the table next to the score chart the risk (%) of persistent symptoms for an individual patient can be determined.

Score chart for prediction of persistent shoulder symptoms at 6 weeks

Duration of complaints		
<6 weeks	0	…
6-12 weeks	7	…
>3 months	11	…
Gradual onset	7	…
Concomitant psychological complaints	10	…
Repetitive movements	8	…
Shoulder pain (0-10)	*score*	…
Neck pain score at PE (0-18)	*score*	…
		+
Total score		…

Total score	**Risk**
≤2	20% - 30%
3 – 7	30% - 40%
8 – 11	40% - 50%
12 – 16	50% - 60%
17– 21	60% - 70%
22 – 27	70% - 80%
28 – 36	80% - 90%
≥37	90% - 100%

PE = Physical Examination

The predicted probability of persistent symptoms at 6 weeks was determined by P=1/[1+ exp – (–1.19 + 0.64 × duration of complaints 6-12 weeks + 0.95 × duration of complaints >3 months + 0.59 × gradual onset + 0.85 × concomitant psychological complaints + 0.68 × repetitive movements + 0.13 × shoulder pain + 0.09 × neck pain score at physical examination)].

Score chart for prediction of persistent shoulder symptoms at 6 months

Duration of complaints		
<6 weeks	0	…
6-12 weeks	9	…
>3 months	17	…
Gradual onset	10	…
Concomitant low back pain	13	…
Shoulder pain (0-10)	*score x 2*	…
Shoulder pain score at PE (0-18)	*score*	…
		+
Total score		…

Total score	**Risk**
≤1	10% - 20%
2 – 16	20% - 30%
17 – 28	30% - 40%
29 – 39	40% - 50%
40– 49	50% - 60%
50 – 61	60% - 70%
≥62	70% - 100%

PE = Physical Examination

The predicted probability of persistent symptoms at 6 months was determined by P=1/[1+ exp – (–1.48 + 0.34 × duration of complaints 6-12 weeks + 0.64 × duration of complaints >3 months + 0.37 × gradual onset + 0.50 × concomitant low back pain + 0.08 × shoulder pain + 0.04 × shoulder pain score at physical examination)].

Figure 3 Prognostic score charts for prediction of persistent symptoms at 6 weeks and 6 months

Discussion

This is the first prospective cohort study on shoulder pain, in which a score chart is developed that may be used by general practitioners to calculate the risk of persistent symptoms for individual patients. Duration of complaints, gradual onset and pain intensity were strong predictors for both short and long term prognosis.

Prognostic factors

In a systematic review[12] of the literature we found only strong evidence for 'high pain intensity' as a predictor of poor outcome. In our study high pain intensity was also found to be a strong predictor of persistent symptoms at short term (6 weeks) and long term (6 months) follow-up. The results of our analyses showed somewhat different sets of predictors for short and long term results, but both analyses demonstrated that duration and severity of symptoms (disease characteristics) were more important in predicting outcome than physical or psychosocial factors. It has previously been suggested that psychosocial factors such as dysfunctional pain cognition or mistaken beliefs about pain and inappropriate pain behaviour are likely to predict a poor outcome of painful musculoskeletal conditions.[5] The association between psychosocial factors and musculoskeletal pain has been established in patients with chronic pain syndromes. The scores on all psychosocial variables measured in our population were low. Although significant univariable associations with persistent symptoms at 6 weeks were found in this study for several psychosocial factors (distress, somatisation, catastrophising), in a multivariable model these factors had little to add to a simple yes or no questions about the presence of psychological complaints. For the applicability of the prediction rules in primary care this is an advantage, as easy to measure predictors are preferred above predictors which are measured with time tasking questionnaires.

Management of shoulder pain

We did not include treatment in the model, as we assumed that confounding by indication could influence our findings. Patients with more severe

symptoms and thus, probably a poorer outcome are more likely to receive more extensive treatment.[37] Only 68 patients (12%) received an injection and 58 (10%) were referred to a physiotherapist, which is a low proportion compared to an earlier study in The Netherlands.[10] The Dutch practice guidelines on shoulder complaints, which recommend a wait-and-see policy during the first 2 to 4 weeks may have lead to a change in practice over the past 5 years. As most patients received wait and see policy or medication, we had a relatively homogeneous group regarding treatment at baseline. Adding treatment variables to our models, indeed, did not improve their predictive value, nor strongly influenced the association of other predictors with outcome (data not shown).

Model fit and discrimination

The calibration plots (Figure 1) show that some predicted probability deciles were slightly too high and some slightly too low. But in general both models are rather well calibrated. The AUCs of the models (0.74 for 6 weeks and 0.67 for 6 months) implied satisfactory discrimination between patients with persistent shoulder symptoms and patients without persistent symptoms.

Analysis

To facilitate comparison between the univariable and multivariable regression analysis we presented uni- and multivariable ORs in Table 1 and 3. In case of high event rates (30-50% risk of persistent symptoms) ORs are an overestimation in comparison to the underlying relative risks (RR), and should not be interpreted as such. In our study we provide, using the prediction rule, the patient and the general practitioner with absolute risks instead of relative risks or odds ratios, because these are easier to understand.

Internal and external validity of collected data

The response to the questionnaires was high (between 83% and 92%)

in this large cohort study. Given the low drop-out rate and only slight differences at baseline between drop-outs and responders we assume that the results can be generalised to all shoulder patients in our study. The GPs were instructed to recruit consecutive patients. We do not have reliable information to gain insight in the percentage of patients who were eligible at first consultation of their GP, and actually participated in the study. In the 10 day period between first consultation and baseline assessment most patients only received advice or medication. Nevertheless, symptoms might have changed in some patients, which may have resulted in a different population at baseline assessment compared to the moment of GP consultation. This could have had an influence on the discriminative ability of our prediction rules. It is difficult to estimate this potential influence as we have no data about changes in symptoms between selection and baseline assessment. This is one of the reasons why we want to stress the importance of validating the prediction rules in a daily practice situation, for which they have been developed.

The recovery rates of 30% after 6 weeks and 54% after 6 months are similar to those found in other studies carried out in primary care populations[9-11], which may strengthen generalisability of our findings to other primary care patients with shoulder pain. However, before considering implementation of our score charts in clinical practice, the generalisability ('external validity') of the models needs to be tested in other populations of patients with shoulder pain.[38] First, the generalisability to another primary care population can be tested. If satisfactory, the generalisability to a community sample, occupational setting, or secondary care population may be tested.

Clinical usefulness

Perhaps most importantly, the clinical usefulness of the developed prediction rules should be established: can the prediction rules be helpful to the clinician when making decisions in the management of patients with shoulder pain, for example, whether or not to consider additional diagnostic testing, start a certain treatment or refer the patient to secondary care.[39] Figure 2 shows that a relatively small proportion of patients is shifted into the lower risk categories at 6 weeks, and a somewhat higher

proportion at 6 months. So, a small number of patients can be reassured by their GP. Patients in the high risk categories possibly benefit from earlier and more extensive treatments. An important objective for future research is to study from which interventions patients in the high risk categories benefit most.

Conclusion

In conclusion, longer duration of symptoms, a gradual onset of symptoms, and high pain intensity at baseline were consistently associated with a poor outcome. The prediction rule and score chart may be used by general practitioners to calculate the absolute risk of persistent symptoms in individual patients with shoulder pain. The performance of our models still needs to be tested in other populations of patients with shoulder pain to enable valid and reliable use of the score charts in clinical practice.

Reference List

1 Bot SD, Van der Waal JM, Terwee CB, Van der Windt DAWM, Schellevis FG, Bouter LM et al. Incidence and prevalence of complaints of the neck and upper extremity in general practice. Ann Rheum Dis 2005;64:118-23.

2 Luime JJ, Koes BW, Hendriksen IJ, Burdorf A, Verhagen AP, Miedema HS et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. Scand J Rheumatol. 2004;33:73-81.

3 Picavet HS, Schouten JS. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3)-study. Pain. 2003;102:167-78.

4 Bongers PM. The cost of shoulder pain at work. BMJ. 2001;322:64-65.

5 Van der Heijden GJMG. Shoulder disorders: a state-of-the-art review. Baillieres Best Pract Res Clin Rheumatol. 1999;13:287-309.

6 Van der Linden MW, Westert GP, De Bakker DH, FG Schellevis. Complaints and disorders in the population and general practice [In Dutch: Klachten en aandoeningen in de bevolking en in de huisartsenpraktijk]. 2005 Utrecht/Bilthoven: NIVEL/RIVM.

7 Okkes IM, Oskam SK, Lamberts H. Van klacht naar diagnose: Episodegegevens uit de huisartspraktijk. Bussum. Coutinho, 1998.

8 Van der Windt DAWM, Koes BWK, Jong BAD, Bouter LM. Shoulder disorders in General Practice: incidence, patients characteristics, and management. Annals of Rheumatic Diseases. 1995;54:959-64.

9 Croft P, Pope D, Silman A. The clinical course of shoulder pain: prospective cohort study in primary care. BMJ. 1996;313:601-2.

10 Van der Windt DAWM, Koes BWK, Boeke AJ, Deville W, De Jong BA, Bouter LM. Shoulder disorders in general practice: prognostic indicators of outcome. Br J Gen Pract. 1996;46:519-23.

11 Winters JC, Sobel JS, Groenier KH, Arendzen JH, Meyboom-de Jong B. The long-term course of shoulder complaints: a prospective study in general practice. Rheumatology. 1999;38:160-3.

12 Kuijpers T, Van der Windt DAWM, Van der Heijden GJGM, Bouter LM. Systematic review of prognostic cohort studies on shoulder disorders. Pain. 2004;109:420-431.

13 Winters JC, Jongh AC, van der Windt DA, Jonquiere M, Winter AF, van der Heijden GJ et al. Practice Guideline "Shoulder complaints' [In Dutch: NHG-Standaard Schouderklachten]. Huisarts en wetenschap. 1999;42:222-31.

14 The Dutch College of General Practitioners (NHG) Practice Guideline. Accessed at http://nhg.artsennet.nl/upload/104/guidelines2/E08.htm on 1 september 2003.

15 Berg SGM, Vlaeyen JWS, Ter Kuil MM, Spinhoven P, Van Breukelen G, Kole-Snijders AMJ. Instruments for measuring chronic pain, part 2. Pain coping and Cognotion List [In Dutch: Meetinstrumenten chronische pijn, deel 2. Pijn Coping en Cognitie Lijst] Maastricht: Pijn Kennis Centrum, 2001.

16 Terluin B. Four-Dimensional Symptom Questionnaire (4DSQ) [In Dutch: De Vierdimensionele Klachtenlijst (4DKL). Een vragenlijst voor het meten van distress, depressie, angst en somatisatie]. Huisarts wet. 1996;39(12):538-47.

17 Terluin B, Van Rhenen W, Schaufeli W, and De Haan M. The four-Dimensional symptom questionnaire (4DSQ): measuring distress and other mental health problems in a working population. Work & Stress 2004;18:187-207.

18 Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain. 1993;52:157-68.

19 Vendrig A, Deutz P, Vink I. Dutch version of the Fear-Avoidance Beliefs Questionnaire [In Dutch: Nederlandse vertaling en bewerking van de Fear-Avoidance Beliefs Questionnaire]. Nederlands Tijdschrift voor Pijn en Pijnbestrijding. 1998;18:11-14.

20 Kori SH, Miller RP, Todd DD. Kinesiophobia: A new view of chronic pain behavior. Pain Mangement 1990;35-43

21 Vlaeyen JWS, Seelen HAM, Peters M, Jong Pd, Aretz E, Beisiegel E et al. Fear Of Movement/(Re)Injury And Muscular Reactivity In Chronic Low Back Pain Patients: An Experimental Investigation. Pain. 1999;82:297-304.

22 Bergman GJD, Van der Heijden GJMG, Winters JC, Groenier KH, Meyboom-de Jong B, Postema K. Manipulative therapy for patients with shoulder complaints: physical examination results in a randomised controlled trial. Submitted, 2004.

23 Neer CS. Impingement lesions. Clin Orthop. 1983;70-77.

24 Solem-Bertoft E, Lundh I, Westerberg CE. Pain is a major determinant of impaired performance in standardized active motor tests. A study in patients with fracture of the proximal humerus. Scand J Rehabil Med. 1996;28:71-78.

25 Westerberg CE, Solem-Bertoft E, Lundh I. The reliability of three active motor tests used in painful shoulder disorders. Presentation of a method of general applicability for the analysis of reliability in the presence of pain. Scand J Rehabil Med. 1996;28:63-70.

26 Van der Windt DAWM, Koes BWK, Deville W, Boeke AJ, De Jong BA, Bouter LM. Effectiveness of corticosteroid injections versus physiotherapy for treatment of painful stiff shoulder in primary care: randomised trial. BMJ. 1998;317:1292-96.

27 Bergman GJD, Winters JC, Groenier KH, Pool JJ, Meyboom-de Jong B, Postema K et al. Manipulative therapy in addition to usual medical care for patients with shoulder dysfunction and pain: a randomized, controlled trial. Ann Intern Med. 2004;141:432-39.

28 Van der Heijden GJMG, Leffers P, Bouter LM. Shoulder disability questionnaire design and responsiveness of a functional status measure. J Clin Epidemiol. 2000;53:29-38.

29 Beurskens AJ, de Vet HC, Koke AJ, Lindeman E, van der Heijden GJ, Regtop W et al. A patient-specific approach for measuring functional status in low back pain. J Manipulative Physiol Ther. 1999;22:144-48.

30 Harrell Jr. 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.

31 Altman DG. Practical statistics for medical research. London: Chapman & Hall, 1991.

32 Efron B, Tibshirani R. An introduction to the bootstrap. Monographs on statistics and applied probability. New York: NY: Chapman and Hall; 1993.

33 Hosmer DW, Lemeshow S. Applied logistic regression. New York: john Wiley & Sons; 1989:135-75.

34 Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules. Applications and methodological standards. N Engl J Med. 1985;313:793-99.

35 Wasson JH, Sox HC. Clinical prediction rules. Have they come of age? JAMA. 1996;275:641-42.

36 Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. JAMA. 1997;277:488-94.

37 Miettinen OS. The need for randomization in the study of intended effects. Stat Med. 1983;2:267-71.

38 Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med. 1999;130:515-24.

39 Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Validity of prognostic models: when is a model clinically useful? Semin Urol Oncol 2002;20:96-107.

5

Generalisability of clinical prediction rules for the prognosis of shoulder pain in general practice

Ton Kuijpers
Geert JMG van der Heijden
Yvonne Vergouwe
Jos WR Twisk
A Joan P Boeke
Lex M Bouter
Danielle AWM van der Windt

ABSTRACT

Background Shoulder pain is common in primary care, and has an unfavorable outcome in many patients. Recently we developed clinical prediction rules for the short and long term prognosis of shoulder pain in general practice.

Objective The objective was to evaluate the generalisability of these prediction rules by applying them to a different but comparable population of patients with a new episode of shoulder pain consulting their general practitioner.

Methods A large research program, consisting of a prognostic cohort study and three randomized controlled trials with 6 months follow-up, was carried out in different geographic area's in The Netherlands. The clinical prediction rules were derived from the results of the prognostic cohort study (n=587). The main outcome measure was persistent symptoms at 6 weeks (short term) or 6 months (long term). The control groups of the trials who received usual care were merged (n=212), and used to validate the prediction rules. Generalisability of the prediction rules was tested by studying calibration and discrimination in the validation cohort.

Results The prediction rule for short term outcome showed reasonable calibration and discriminative ability in the validation cohort. The area under the ROC curve (AUC) was 0.72 compared to 0.74 in the derivation cohort. The prediction rule for long term outcome performed less well. Discriminative ability (AUC) decreased to 0.56 in the validation cohort compared to 0.67 in the derivation cohort.

Conclusions The prediction rule for the short term prognosis of shoulder pain in general practice showed good generalisability. The prediction rule for the long term prognosis showed poor generalisability. Hence, it seems difficult to make accurate predictions of the long term outcome of shoulder pain in general practice.

Introduction

Shoulder pain is common with a one year prevalence ranging between 5% and 47%.[1-5] The prevalence in the general population in The Netherlands has recently been estimated at 17%.[6] The annual incidence of consultation for a new episode of shoulder pain in Dutch general practice ranges between 12 and 25/1000/year.[3;6-8] Shoulder pain has an unfavorable outcome in many patients. About 40 to 50% of all patients who present with a new episode of shoulder pain in primary care report persistent symptoms after 6 to 12 months.[9-11]

We developed clinical prediction rules consisting of a limited number of (easily measurable) prognostic factors to predict the risk of persistent shoulder symptoms at the short (6 weeks) and long term (6 months). Such information may also support decisions regarding treatment and referral of patients. The performance (that is, calibration and discrimination) of the prediction rules was evaluated in the development study[12]. Calibration refers to what extent the observed frequencies agree with the predicted probabilities. Discrimination refers to the ability to distinguish between a patient with persistent symptoms and a patient without persistent symptoms.

Before considering implementation of the prediction rules in clinical practice their generalisability needs to be tested.[13-15] Generalisability refers to the performance in patients drawn from a different but comparable population.[13] Our objective was therefore to evaluate the performance of our clinical prediction rules for the prognosis of shoulder pain in a different population of patients with shoulder pain in primary care.

Methods

Dutch Shoulder Study

The Dutch Shoulder Study (DSS) is a comprehensive cohort study, carried out between January 2000 and May 2005. The DSS consists of one prognostic cohort study and three randomised controlled trials, which were carried out alongside each other. Between January 2001 and June 2003, 103 general practitioners (GP) recruited patients at first

Table 1 Selection criteria of the Dutch Shoulder Study

General inclusion criteria
Patients older than 18 years of age Not consulted GP or received any form of treatment for the afflicted shoulder in the preceding 3 months Sufficient knowledge of the Dutch language
Specific inclusion criteria trials
Groningen Manipulation Study (GMO) Dysfunction of the cervicothoracic spine and adjacent ribs with accompanying pain or restricted movement
Graded Exercise Therapy Study (GET) Duration of complaints >3 months
Education and Activation Program (EAP) Duration of complaints <3 months
Exclusion criteria
Severe physical or psychological conditions (i.e. fractures or luxation in the shoulder region; rheumatic disease; neoplasm; neurological or vascular disorders; dementia)

consultation for a new episode of shoulder complaints in three geographic areas in the Netherlands (Amsterdam, Groningen and Maastricht). All patients in the DSS had to meet the same general inclusion criteria, and specific additional inclusion criteria if eligible for a trial (Table 1). For the prognostic cohort study no additional inclusion criteria were specified. Data of the prognostic cohort study were used to derive the prediction rules. Data of the control groups of the three trials were used to study the generalisability of the rules.

The Groningen Manipulation Study (GMO)[16;17] studies the effectiveness of manipulative therapy for the shoulder girdle in addition to usual care. In two other trials a Graded Exercise Therapy (GET)[18] and an Education and Activation Program (EAP)[19], respectively, were studied. Patients in the control groups of the trials received usual care, similar to the patients in the cohort study.

Baseline and follow-up assessments for all patients in the DSS were identical. The outcome was measured by postal questionnaires at 6 weeks, 3 and 6 months. The primary outcome measure 'Patient perceived recovery' was measured on an 8-point scale. Patients who did not report full recovery or very much improvement were denoted as having "persistent symptoms".[17;20] Secondary outcome measures were shoulder disability, measured with the 16-item shoulder disability questionnaire

(SDQ; 0-100)[21], pain (0-10 numeric rating scale)[20], and severity of the main complaint (0-10 numeric rating scale)[22].

Management of shoulder pain

All participants in the cohort study and the trial participants randomised to the control groups, received standardised treatment according to the 1999 version of the Dutch guidelines for shoulder disorders issued by the Dutch College of General Practitioners.[23;24] The guidelines recommend giving information on the prognosis of shoulder pain, advice regarding provoking activities, and stepwise treatment consisting of paracetamol, Non-Steroidal Anti-Inflammatory Drugs (NSAIDs), corticosteroid injection or referral for physiotherapy. The GP made the decision regarding the content of treatment based on duration and severity of pain and disability. The participating general practitioners were educated and trained to apply treatment according to this guideline.

Prediction rules

The prediction rules for persistent symptoms (yes/no) after 6 weeks and 6 months were developed using information from the 587 patients of the derivation cohort. Sociodemographic variables, disease characteristics (i.e. pain intensity, disability, duration of complaints, onset, comorbidity), physical workload, psychosocial factors and results of a physical examination were used to derive the prediction rules. We tested the internal validity with bootstrapping techniques.[13] The calibration of the prediction rules was adequate. The discriminative ability was satisfactory with area under the receiver operating characteristic (ROC) curve of 0.74 (95% CI 0.70; 0.79) at 6 weeks and 0.67 (95% CI 0.63; 0.71) at 6 months. Figure 1 presents the prediction rules as score charts, which development has been described in detail elsewhere.[12]

Instruction
If a predictor is scored positively, the given weight needs to be filled in. Subsequently the scores are added to calculate the 'Total score'. From the table next to the score chart the risk (%) of persistent symptoms for an individual patient can be determined.

Score chart for prediction of persistent shoulder symptoms at 6 weeks

Duration of complaints		
<6 weeks	0	…
6-12 weeks	7	…
>3 months	11	…
Gradual onset	7	…
Concomitant psychological complaints	10	…
Repetitive movements	8	…
Shoulder pain (0-10)	*score*	…
Neck pain score at PE (0-18)	*score*	…
		+
Total score		…

Total score	Risk
≤2	20% - 30%
3 – 7	30% - 40%
8 – 11	40% - 50%
12 – 16	50% - 60%
17– 21	60% - 70%
22 – 27	70% - 80%
28 – 36	80% - 90%
≥37	90% - 100%

PE = Physical Examination

The predicted probability of persistent symptoms at 6 weeks was determined by P=1/[1+ exp – (–1.19 + 0.64 × duration of complaints 6-12 weeks + 0.95 × duration of complaints >3 months + 0.59 × gradual onset + 0.85 × concomitant psychological complaints + 0.68 × repetitive movements + 0.13 × shoulder pain + 0.09 × neck pain score at physical examination)].

Score chart for prediction of persistent shoulder symptoms at 6 months

Duration of complaints		
<6 weeks	0	…
6-12 weeks	9	…
>3 months	17	…
Gradual onset	10	…
Concomitant low back pain	13	…
Shoulder pain (0-10)	*score x 2*	…
Shoulder pain score at PE (0-18)	*score*	…
		+
Total score		…

Total score	Risk
≤1	10% - 20%
2 – 16	20% - 30%
17 – 28	30% - 40%
29 – 39	40% - 50%
40– 49	50% - 60%
50 – 61	60% - 70%
≥62	70% - 100%

PE = Physical Examination

The predicted probability of persistent symptoms at 6 months was determined by P=1/[1+ exp – (–1.48 + 0.34 × duration of complaints 6-12 weeks + 0.64 × duration of complaints >3 months + 0.37 × gradual onset + 0.50 × concomitant low back pain + 0.08 × shoulder pain + 0.04 × shoulder pain score at physical examination)].

Figure 1 Prognostic score charts for prediction of persistent symptoms at 6 weeks and 6 months

Analysis

The performance of the prediction rules was tested in the validation cohort by evaluating their calibration and discrimination. Calibration was assessed by plotting the predicted probabilities of persistent symptoms according to the prediction rule, against the observed frequencies. For this, patients were grouped into quintiles according to their predicted probability of persistent symptoms. The prevalence of the endpoint within each quintile equals the observed frequency. A more formal indication of calibration can be obtained by fitting a logistic regression model with the logodds of the predicted risks as only covariate. This model has an intercept and a slope. If predicted risks and observed frequencies are in agreement, the intercept is equal to 0 and the slope equal to 1.

The area under the ROC curve was used to assess the discriminative ability of the model. An area under the curve (AUC) of 0.5 indicates no discrimination above chance, whereas an AUC of 1.0 indicates perfect discrimination. Since the discriminative ability of a rule is related to the homogeneity of the sample in which the rule is applied, we estimated the maximum attainable AUC. Using the predicted risks of the patients in the validation cohort, outcomes were generated with Monte Carlo Simulation.[25;26] This mimics the situation that the model is perfectly calibrated. The AUC that was estimated for the predicted risks and generated outcomes was considered the maximum attainable AUC for the validation sample.

Furthermore, to gain insight in the performance of our prediction rules, we estimated the multivariable logistic regression coefficients for each of the predictors of our prediction rule in the validation cohort. This analysis shows which of the different elements of the rule are the strongest predictors of persistent shoulder pain in the validation cohort.

Results

Study population

Table 2 presents the baseline characteristics of the derivation cohort and validation cohort. Patients in the validation cohort clearly showed a longer

Table 2 Description of baseline characteristics and outcome measures at 6 weeks and 6 months of patients with shoulder pain in the derivation (n=587) and validation cohort (n=212)

Baseline characteristics	Derivation		Validation	
Demographic				
Age (years); mean (SD)	51 (14)		51 (12)	
Gender: male; n (%)	292 (50)		92 (44)	
Disease characteristics				
Duration of complaints; n (%)	205 (35)		46 (22)	
0-6 weeks*	139 (24)		49 (23)	
7-12 weeks	242 (41)		115 (55)	
>3 months	363 (62)		144 (69)	
Gradual onset (vs. acute); n (%)				
Precipitating cause; n (%)	138 (24)		58 (28)	
Strain/overuse: usual activities	348 (62)		136 (65)	
Shoulder complaints in the past; n (%)	296 (51)		128 (61)	
Neck complaints in the past; n (%)	55 (9)		20 (10)	
Comorbid psychological complaints; n (%)				
Concomitant musculoskeletal complaints; n (%)				
Neck/high back	209 (36)		85 (41)	
Low back pain	139 (24)		59 (28)	
Upper extremity	174 (30)		76 (36)	
Shoulder pain (0-10); mean (SD)	4.8 (2.3)		5.3 (2.2)	
Shoulder disability (0-100); mean (SD)	60 (24)		62 (24)	
Physical examination				
Shoulder pain score (0-18); median (IQR)	4 (2-4)		7 (4-7)	
Neck pain score (0-18); median (IQR)	0 (0-0)		2 (0-2)	
Physical factors				
Dynamic physical workload (0-5); median (IQR)	1 (1-2)		1 (0-1)	
Repetitive movements; n (%)	384 (65)		151 (73)	
Outcome measures	**6 weeks**	**6 months**	**6 weeks**	**6 months**
Persistent symptoms; n (%)	340 (70)	249 (46)	161 (89)	125 (69)
Pain* (0-10); mean (SD)	4.3 (2.1)	4.1(2.3)	4.3 (2.1)	4.0 (2.0)
Shoulder disability* (0-100); mean (SD)	53.0 (25.5)	52.2 (26.7)	56.0 (25.6)	54.4 (27.2)
Severity of main complaint* (0-10); mean (SD)	4.8 (2.6)	5.0 (2.8)	4.9 (2.5)	5.6 (2.6)

SD= standard deviation; IQR = Inter quartile range; *Mean and SD presented for group reporting persistent symptoms.

duration of complaints at baseline (13% less often complaints between 0-6 weeks, and 14% more often complaints >3 months), and reported 10% more neck complaints in the past in comparison with the derivation cohort.

Course of symptoms

Table 2 shows that more patients in the validation cohort reported persistent symptoms after 6 weeks (89% versus 70%) and 6 months (69% versus 46%), compared to the derivation cohort. Nevertheless, patients in the derivation and validation cohort reported similar levels of pain and disability at the different time points. Likewise, patients reporting recovery in the validation cohort and in the derivation cohort showed similar low levels of pain (<1 point), disability (<13 points), and severity of the main complaint (<1 point).

Management of shoulder pain

At baseline most patients in the derivation cohort (n=423, 72%) received a wait and see policy, paracetamol, or NSAIDs. Furthermore, 68 patients (12%) received an injection with corticosteroid, 58 patients (10%) were referred for physiotherapy and 28 patients (6%) received other therapies. In the validation cohort 141 patients (83%) received a wait and see policy, paracetamol, or NSAIDS; 9 patients (5%) received an injection; 11 patients (7%) were referred for physiotherapy and 8 patients (5%) received other therapies.

Performance

Figure 2 shows the calibration of the predictions. For 6 weeks the plotted points were rather close to the 45° line, although most predictions slightly underestimated the observed probabilities. This is in agreement with an intercept of 0.44 in a model, using the logodds of the predicted probabilities as only covariate. The intercept above 0 confrims that the predicted probabilities were generally too low. The slope of the model was 1.1, which is close to 1.

For 6 months the plotted points were further away of the 45° line, demonstrating a rather poor calibration of the model. This is confirmed by an intercept of 0.39 (not close to 0) and a slope of 0.43 (not close to 1).

The discriminative ability (AUCs) of the prediction rules was 0.72 (95%

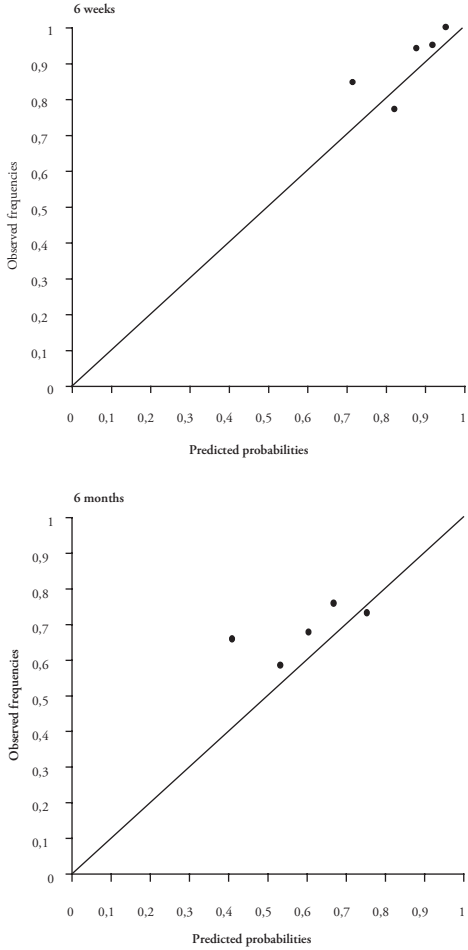


Figure 2 Calibration plots showing the observed frequencies versus the predicted probabilities for persistent symptoms at 6 weeks (n=175) and 6 months (n=180) in the validation cohort.

Table 3 Prediction rules for persistent shoulder symptoms at 6 weeks and 6 months after first consultation.

		Derivation		Validation	
Predictor	**Scale**	β	95% CI	β	95% CI
6 weeks					
Duration of complaints					
0-6 weeks*		-		-	
7-12 weeks	(yes/no)	0.64	0.1, 1.2	0.42	-0.9, 1.7
>3 months	(yes/no)	0.95	0.4, 1.5	1.67	0.2, 3.1
Gradual onset	(yes/no)	0.59	0.1, 1.1	0.36	-0.7, 1.4
Concomitant psychological complaints	(yes/no)	0.85	-0.1, 1.9	0.74	-1.5, 3.0
Repetitive movements	(yes/no)	0.68	0.2, 1.1	1.05	-0.1, 2.2
Shoulder pain	(0-10)	0.13	0.0, 0.2	0.06	-0.2, 0.3
Neck pain score at physical examination	(0-18)	0.09	0.0, 1.0	0.03	-0.1, 0.2
6 months					
Duration of complaints					
0-6 weeks*		-		-	
7-12 weeks	(yes/no)	0.34	-0.1, 0.8	-0.10	-1.3, 1.1
>3 months	(yes/no)	0.64	0.2, 0.1	1.24	-0.1, 2.6
Gradual onset	(yes/no)	0.37	0.0, 0.6	0.70	-0.3, 1.7
Concomitant low back pain	(yes/no)	0.50	0.1, 0.9	0.01	-1.1, 1.2
Shoulder pain	(0-10)	0.08	0.0, 0.2	0.04	-0.2, 0.3
Shoulder pain score at physical examination	(0-18)	0.04	0.0, 0.1	0.09	-0.0, 0.2

The β's are derived from a multiple logistic regression analysis. The β's for the validation cohort are derived from the results of a multiple logistic regression analysis conducted with the predictors from the derivation cohort. *Reference category CI=Confidence Interval

CI 0.63; 0.82) at 6 weeks, and 0.57 (95% CI 0.48; 0.66) at 6 months. The Monte Carlo Simulation showed a maximum attainable AUC of 0.70 at 6 weeks, and 0.64 at 6 months.

Table 3 shows the estimates of the regression coefficients for each predictor when the rule is applied to the validation cohort. A duration of complaints >3 months and repetitive movements were strong predictors of persistent symptoms at 6 weeks. For long term outcomes a duration >3 months and a gradual onset seemed strongly related with persistent symptoms at 6 months.

Discussion

The performance of the prediction rule for the short term (6 weeks) prognosis of shoulder pain in the validation cohort was satisfactory. Calibration and discriminative ability were reasonable, and similar to

that found in the derivation cohort (AUC=0.74 in derivation cohort, and AUC=0.72 in validation cohort). The prediction rule for long term (6 months) outcome showed poor calibration and discrimination. The AUC decreased from 0.67 in the derivation cohort to 0.57 in the validation cohort, not much more than a flip of a coin (AUC=0.50), which means that the performance of the long term prediction rule was disappointing.

The patients in the validation cohort differed from the derivation cohort regarding several aspects. In general, the shoulder complaints from the patients in the validation cohort were more severe (Table 2). They showed a longer duration of symptoms, which is an important predictor of outcome[12], and reported more neck complaints in the past. This may indicate that we have tested the performance of the prediction rule in patients who, on average, were more advanced in their disease process or who may have had a somewhat different type of shoulder problem (which has been described as spectrum transportability[13]). More severe complaints at baseline may have resulted in more frequent reports of persistent symptoms (Table 2). Another possible explanation for the higher occurrence of persistent symptoms could be that the validation cohort more often received a wait and see policy (83% versus 72%) and were less frequently treated with local infiltration of a corticosteroid (5% versus 12%).

Differences in prognosis between the derivation and validation cohort may have substantially altered the calibration of our prediction rules in the validation cohort, especially for the long term. The reason for this is that statistical models are calibrated to the overall outcome prevalence. For a substantial part, this prevalence is determined by the characteristics of the patient population. As long as the overall prevalence is explained by predictors which are included in the prediction rule, the model will still be well calibrated. This may have resulted in a reasonable calibration of our short term prediction rule, despite differences at baseline between the derivation and validation cohort regarding important predictors, i.e. duration of complaints and repetitive movement. The poor performance of the prediction rule for long term outcomes may be explained by patient characteristics which are not documented in this study, but yet strongly influenced outcome in the validation cohort.

The maximum attainable AUC for the short term of 0.70 strengthens our findings of adequate discriminative ability of this rule. Calibration was reasonable, although an intercept of 0.44 showed that the risk of persistent shoulder pain was generally underestimated (see also Figure 2). For the long term, a substantially lower maximum attainable AUC of 0.64 differed considerably from the achieved AUC of 0.57. This indicated a model with poor discriminative ability. Regression coefficients were generally to high, and insufficient shrinking of the regression coefficients had been achieved in the development stage of the prediction rule. Justice[13] stated that perhaps the most difficult test of discrimination occurs when the spectrum of a disease narrows from both sides; that is, the test sample includes many patients who have an illness of intermediate severity and very few who are either severely ill or not very ill at all. This could partly explain the poor performance of our long term prediction rule as most observed probabilities of persistent symptoms were distributed between 0.5 and 0.7 (Figure 2). This reflects a homogeneous population resulting in a low maximum attainable AUC of 0.64.

We developed prediction rules to predict the prognosis of shoulder pain in general practice. Most elements of the prediction rules were derived from a questionnaire, filled out by the patient. If these prediction rules would be implemented in daily practice it is the general practitioner who asks the questions and calculates the risk by using a score chart. Or in a more sophisticated way, enters the responses into a personal computer (PC) or personal digital assistant (PDA), which calculates the risk of persistent symptoms. So, future research should also evaluate the methodologic transportability of the prediction rules (i.e. performance when data are collected by using alternative methods[13]) in a new sample of patients. And perhaps most importantly, the clinical usefulness of these instruments should be established: can the prediction rules be helpful to the clinician when making decisions in the management of patients with shoulder pain, for example, whether or not to consider additional diagnostic testing, start a certain treatment or refer the patient to secondary care.[15]

In conclusion, the prediction rule for the short term (6 weeks) prognosis of shoulder pain in general practice showed adequate generalisability in

the validation cohort. For the long term outcome (6 months) it seems difficult to make accurate predictions of persistent shoulder pain in this population.

Reference List

1. Luime JJ, Koes BW, Hendriksen IJ, et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. Scand J Rheumatol 2004;33:73-81.
2. Van der Heijden GJMG. Shoulder disorders: a state-of-the-art review. Baillieres Best Pract Res Clin Rheumatol 1999;13:287-309.
3. Bot SD, Van der Waal JM, Terwee CB, Van der Windt DAWM, Schellevis FG, Bouter LM et al. Incidence and prevalence of complaints of the neck and upper extremity in general practice. Ann Rheum Dis 2005;64:118-23.
4. Bongers PM. The cost of shoulder pain at work. BMJ. 2001;322:64-65.
5. Picavet HS, Schouten JS. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3)-study. Pain. 2003;102:167-78.
6. Van der Linden MW, Westert GP, De Bakker DH, FG Schellevis. Complaints and disorders in the population and general practice [In Dutch: Klachten en aandoeningen in de bevolking en in de huisartsenpraktijk]. 2005 Utrecht/Bilthoven: NIVEL/RIVM.
7. Van der Windt DAWM, Koes BWK, Jong BAD, Bouter LM. Shoulder disorders in General Practice: incidence, patients characteristics, and management. Annals of Rheumatic Diseases. 1995;54:959-64.
8. Okkes IM, Oskam SK, Lamberts H. Van klacht naar diagnose: Episodegegevens uit de huisartspraktijk. Bussum. Coutinho, 1998.
9. Croft P, Pope D, Silman A. The clinical course of shoulder pain: prospective cohort study in primary care. BMJ 1996;313:601-2.
10. Van der Windt DAWM, Koes BWK, Boeke AJP, et al. Shoulder disorders in general practice: prognostic indicators of outcome. Br J Gen Pract 1996;46:519-23.
11. Winters JC, Sobel JS, Groenier KH, et al. The long-term course of shoulder complaints: a prospective study in general practice. Rheumatology 1999;38:160-3.
12. Kuijpers T, Van der Windt DAWM, Van der Heijden GJMG, Twisk J, Vergouwe Y, Bouter LM. Clinincal prediction rules for the prognosis of shoulder pain in general practice. Pain, in press.
13. Justice AC, Covinsky KE, Berlin JA. Assessing the generalisability of prognostic information. Ann Intern Med 1999;130:515-24.
14. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000;19:453-73.
15. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Validity of prognostic models: when is a model clinically useful? Semin Urol Oncol 2002;20:96-107.
16. Bergman GJ, Winters JC, Van der Heijden GJMG, Postema K, Meyboom-de Jong B. Groningen Manipulation Study. The effect of manipulation of the structures of the shoulder girdle as additional treatment for symptom relief and for prevention of chronicity

or recurrence of shoulder symptoms. Design of a randomized controlled trial within a comprehensive prognostic cohort study. J Manipulative Physiol Ther 2002;25:543-9.

17. Bergman GJ, Winters JC, Groenier KH, Pool JJ, Meyboom-de Jong B, Postema K et al. Manipulative therapy in addition to usual medical care for patients with shoulder dysfunction and pain: a randomized, controlled trial. Ann Intern Med 2004;141:432-9.
18. Geraets JJ, Goossens ME, De Bruijn CP, Koke AJ, De Bie RA, Pelt RA et al. A behavioural treatment for chronic shoulder complaints: concepts, development, and study design. Aust J Physiother 2004;50:33-8.
19. De Bruijn CP, De Bie R, Geraets JJ, Goossens ME, Koke AL, Van den Heuvel HW et al. Evaluation of an education and activation programme to prevent chronic shoulder complaints: design of an RCT [ISRCTN71777817]. BMC Fam Pract 2005;6:7.
20. Van der Windt DAWM, Koes BWK, Deville W, Boeke AJ, De Jong BA, Bouter LM. Effectiveness of corticosteroid injections versus physiotherapy for treatment of painful stiff shoulder in primary care: randomised trial. BMJ. 1998;317:1292-96.
21. Van der Heijden GJMG, Leffers P, Bouter LM. Shoulder disability questionnaire design and responsiveness of a functional status measure. J Clin Epidemiol. 2000;53:29-38.
22. Beurskens AJ, de Vet HC, Koke AJ, Lindeman E, van der Heijden GJ, Regtop W et al. A patient-specific approach for measuring functional status in low back pain. J Manipulative Physiol Ther. 1999;22:144-48.
23. The Dutch College of General Practitioners (NHG) Practice Guideline. Accessed at http://nhg.artsennet.nl/upload/104/guidelines2/E08.htm on 1 september 2003.
24. Winters JC, Jongh AC, van der Windt DA, Jonquiere M, Winter AF, van der Heijden GJ et al. Practice Guideline "Shoulder complaints' [In Dutch: NHG-Standaard Schouderklachten]. Huisarts en wetenschap. 1999;42:222-31.
25. Gilks WR RSSD. Markov Chain Monte Carlo in practice. New York: Chapman & Hall, 1996.
26. Vergouwe Y, Steyerberg EW, Eijkemans M, Habbema J. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. J Clin Epidemiol, in press. 2005

6

A prediction rule for shoulder pain related sick leave

Ton Kuijpers
Daniëlle AWM van der Windt
Geert JMG van der Heijden
Jos WR Twisk
Yvonne Vergouwe
Lex M Bouter

ABSTRACT

Background Shoulder pain is common in primary care, and has an unfavourable outcome in many patients. Information about predictors of shoulder pain related sick leave in workers is scarce and inconsistent.

Objective To develop a clinical prediction rule for calculating the risk of shoulder pain related sick leave for individual workers, during the 6 months following first consultation in general practice.

Methods A prospective cohort study with 6 months follow-up was conducted among 350 workers with a new episode of shoulder pain. Potential predictors included the results of a physical examination, sociodemographic variables, disease characteristics (duration of symptoms, sick leave in the 2 months prior to consultation, pain intensity, disability, comorbidity), physical activity, physical workload, and (work related) psychosocial factors. The main outcome measure was sick leave during 6 months following first consultation in general practice.

Results Response rate to the follow-up questionnaire at 6 months was 85%. During the 6 months after first consultation 30% (89/298) of the workers reported sick leave. 16% (47) reported 10 days sick leave or more. Sick leave during this period was predicted in a multivariable model by a longer duration of sick leave prior to consultation, more shoulder pain, strain (overuse) as a result of usual activities, and concomitant psychological complaints. The discriminative ability of the prediction model was satisfactory with an area under the curve of 0.70 (95% CI 0.64; 0.76).

Conclusions Although 30% of all workers with shoulder pain reported sick leave during follow-up, the duration of sick leave was limited to a few days in most workers. We developed a prediction rule and a score chart that can be used by general practitioners and occupational health care providers to calculate the absolute risk of sick leave in individual workers with shoulder pain, which may help to identify workers who need additional attention. The performance of our models needs to be tested in other working populations with shoulder pain to enable valid and reliable use of the score charts in everyday practice.

Introduction

Shoulder pain is common with a one-year prevalence ranging between 5% and 47%.[1-7] In occupational settings, the one year prevalence ranges between 16% and 47%.[3-5] A Finnish study[8] reported a one-year incidence of shoulder pain of 14% among forestry workers. Shoulder pain has an unfavourable outcome in many patients. About 50% of all new episodes of shoulder disorders presented in primary care show complete recovery within 6 months[9-11], while after one year this proportion increases to only 60%.[10] In an occupational study a one year persistence of shoulder pain of 55% was reported.[8]

In a systematic review of the literature we summarized the available evidence from 16 studies regarding prognostic factors of shoulder disorders.[12] Only six studies were of relatively high quality. In an occupational setting strong evidence for predicting poorer outcome was only found for age (45-54 years). Evidence for (work related) psychosocial variables was lacking.[2] In a systematic review of occupational risk factors it has been suggested that many factors play a role in the occurrence of new episodes of shoulder pain, including physical load and the psychosocial work environment.[13] Also in studies on other musculoskeletal disorders, such as low back pain or neck pain, a relationship between work-related psychosocial factors and the occurrence of pain or sickness absence has been reported.[14;15] Possibly these factors are also of importance in the persistence of shoulder pain related sick leave, and may help to identify workers who need additional attention.

We performed a cohort study among workers who had presented shoulder complaints to their general practitioner, and followed them for 6 months. Our objective was to develop a score chart to identify workers who will report at least 1 day of shoulder pain related sick leave during 6 months following first consultation for their complaints.

Methods

Study population

Between January 2001 and June 2003, 103 general practitioners recruited

workers with a new episode of shoulder disorders in three geographic area's in the Netherlands (Amsterdam, Groningen and Maastricht). Workers were selected if they were older than 18 years of age, had a paid job, and did not consult their GP or receive any form of treatment in the preceding 3 months for the afflicted shoulder. Sufficient knowledge of the Dutch language was required to complete written questionnaires. Exclusion criteria were severe physical or psychological conditions (i.e. fractures or dislocation in the shoulder region; rheumatic disease; neoplasm; neurological or vascular disorders; dementia). There was no restriction with respect to type of work or occupation.

Management of shoulder disorders

All workers received standardised treatment according to the 1999 version of the Dutch guidelines for shoulder disorders issued by the Dutch College of General Practitioners[16;17] which consists of information on the prognosis of shoulder pain, advice regarding provoking activities, and stepwise treatment consisting of paracetamol, NSAIDs, corticosteroid injection or referral for physiotherapy. The GP made the decision regarding the content of treatment based on duration and severity of pain and disability. The participating general practitioners were educated and trained to apply treatment according to this guideline.

Prognostic factors

Within a few days after consultation all workers completed a questionnaire, which contained questions on sociodemographic variables, disease characteristics (i.e. pain intensity, disability, duration of complaints, sick leave in the 2 months prior to consultation, pain onset, comorbidity), physical activity, physical workload, work-related psychosocial factors, and individual psychological factors. A physical examination was carried out by a trained assistant.

Physical activity was measured with a single question (less/equally/more active than others). We measured physical workload with a self-constructed scale of 5 questions concerning pushing and pulling, lifting weights, and

working with hands above shoulder level, on at least two days a week (total score 0-5, Crohnbach's α= 0.74). Repetitive movements and sitting in the same position for extended time periods, on at least two days a week, were measured with single questions answered by yes or no.

Work-related psychosocial factors, were assessed with the 27-item Job Content Questionnaire (JCQ)[18], which measures all dimensions of the widely used Demand-Control-Support model. On a four point scale (totally disagree, disagree, agree, totally agree) workers rated certain aspects of their work. The JCQ consists of the dimensions quantitative job demands (4-20); skill discretion (4-20); decision authority (4-12); co-worker support (4-16) and supervisor support (4-16), as proposed by Karasek et al.[18] and clinimetrically evaluated by De Jonge et al.[19]

The following individual psychosocial variables were measured: coping, anxiety, depression, somatization, distress, fear-avoidance beliefs, and kinesiophobia. Coping was assessed with the 43-item Pain Coping and Cognition List (PCCL)[20], consisting of the subdomains catastrophizing (1-6 points), coping with pain (1-6 points), internal (1-6 points) and external locus of control (1-6 points). Anxiety (0-24 points), depression (0-12 points), somatization (0-32 points), and distress (0-32 points), were measured with the 50-item Four-Dimensional Symptom Questionnaire (4DSQ).[21] Fear-avoidance beliefs were assessed using the 4-item physical activity subscale of the Fear Avoidance Beliefs Questionnaire (FABQ; 0-24).[22] Kinesiophobia was measured using two items of the Tampa Scale for Kinesiophobia (TSK; 0-12).[23] Our baseline questionnaire also included a general one-item question regarding the presence (yes/no) of any psychological problems (e.g. distress, depression, anxiety).

The physical examination by a research assistant contained testing of the shoulder joint and cervicothoracic spine. For the glenohumeral joint active and passive abduction, passive exorotation, and shoulder impingement[24] were tested. Two alternative functional tests, HIB (Hand-in-back) and HIN (Hand-in-neck)[25;26] measured on a 7-point scale were performed as well. The mobility of the cervicothoracic spine was tested with flexion, with extension, with rotation in flexed, extended and neutral position and with lateral bending. During all mobility tests pain was assessed on a 4-

point scale (0 = no pain; 3 = severe pain). A factor analysis on the results of a physical examination in a similar population of patients with shoulder disorders resulted in four factors: shoulder mobility (0-18), shoulder pain (0-18), neck mobility (0-4) and neck pain (0-18).[27]

Outcome measurements

The outcome was measured by postal questionnaires at 6 weeks, 3 and 6 months. Our primary outcome measure was sick leave due to shoulder pain (yes =≥1 day, no = 0 days). Secondary outcome measures were patient perceived recovery, shoulder disability, measured with the 16-item shoulder disability questionnaire (SDQ; 0-100)[28], shoulder pain (0-10 numeric rating scale)[29], and severity of the main complaint (0-10 numeric rating scale)[30]. We studied the relationship between our primary and secondary outcome measures to determine if workers reporting sick leave during follow-up showed higher levels of pain and disability.

Analysis

Missing values of patient characteristics were imputed (approx. 2% of all required values). Imputation was based on the correlation between the variable with missing values with the other patient characteristics. Univariable logistic regression analyses were performed for all potential prognostic indicators with our primary outcome measure, i.e. sick leave during 6 months following first consultation. Variables that had a statistically significant association with the outcome (p-value ≤0.20) were selected for the backward selection in the multivariable analysis, and checked for co-linearity. If the correlation between the determinance was higher than 0.5, the most feasible was included in our analysis. We adopted a hierarchically approach in the variable selection in which easily obtainable predictors were included first. Therefore, variables were selected in blocks of increasing effort to obtain during consultation: 1) socio-demographic factors and disease characteristics; 2) physical factors; 3) work-related psychosocial factors; 4) individual psychological factors; 5) physical examination. Variables with the lowest predictive value were

deleted from the model until further elimination of a variable resulted in a statistically significant lower model fit estimated with the likelihood ratio test ($p \leq 0.20$).

Prediction models, usually provide too extreme estimates, when in the development phase no correction is applied. Therefore, we used bootstrap samples to estimate a shrinkage factor.[31] Bootstrap samples were drawn with replacement (100 replications) from the full data set. The backward selection of variables and model fitting was repeated within each bootstrap sample.

Bootstrapping techniques were also used to study the internal validity of the final prediction model.[31;32] The model's performance obtained after bootstrapping can be considered as the performance that can be expected in similar future patients. All analyses were performed using S-plus 6.1 (Insightful Corp., Seattle, WA, USA).

Evaluation of the model

The reliability of the multivariable model was determined with the Hosmer-Lemeshow goodness-of-fit statistic.[33] Calibration of the model predictions, which is related to reliability, was assessed by plotting the predicted individual probability against the observed sick leave. For this, workers were grouped into quintiles according to their predicted probability for sick leave according to the model. The prevalence of the endpoint within each quintile represents the observed frequency. The area under the receiver-operating characteristic curve (ROC) was used to assess the performance of the model in terms of accuracy of correct prediction. The ROC-curve is a plot of the true positive rate (sensitivity) against the false-positive rate (1-specificity) of the model. The curve illustrates the ability of the model to discriminate between workers with and without sick leave at subsequent cut-off points along the range of the predicted probabilities. An area under the curve (AUC) of 0.5 indicates no discrimination above chance, whereas an AUC of 1.0 indicates perfect discrimination.

Prediction of an individual patient's risk

We developed a clinical prediction rule[34-36] for sick leave during 6 months following first consultation, to provide general practitioners and occupational health care providers with an estimate of the absolute risk of sick leave for individual workers. Since we used logistic regression, the probability (P) of sick leave was predicted with $P=1/[1+ \exp-(a_0 + b_1x_1 + \ldots\ldots + b_jx_j)]$. The status of a patient for any dummy or binary variable included in the prediction rule can be either 0 or 1, while for a (semi) continuous variable it takes the actual observed value.

Score chart

To facilitate the calculation of an individual worker's risk, we developed a score chart. We multiplied the regression coefficients by 4 and rounded them to the nearest integer to form the scores for each of the predictors. The sum of these scores correspond to a range of risks of sick leave during follow-up.

Results

Study population and follow-up

A total of 350 workers with shoulder pain in primary care completed the baseline assessment. Table 1 lists the baseline characteristics. At 6 months 298 (85%) workers returned the postal questionnaire.

The drop-outs at 6 months (n=52) showed significantly ($p<0.10$) more pain of the shoulder (2.4 vs. 2.3 points) and the neck (2.8 vs. 2.6 points) at physical examination and less decision authority (4.4 vs. 5.5 points) at baseline.

During the 6 months following first consultation 30% (89/298) of the workers reported at least one day of sick leave (primary outcome measure) because of their shoulder pain, and 16% (47/298) of the workers reported more than 10 days of shoulder pain related sick leave in 6 months. For 25 of the 89 workers with sick leave, this was limited to the first 6 weeks of follow-up. Table 2 shows that workers with sick leave reported also more

Table 1 Baseline characteristics of a working population with shoulder disorders (n=350), and univariable associations with sick leave (yes/no) during 6 months following first consultation in general practice

Variable	n (%)	OR	95% CI	p[a]
Demographic				
Age (years); mean (SD)	45 (11)	1.0	1.0, 1.0	0.26
Gender: male	193 (55)	0.9	0.5, 1.4	0.56
Education				0.02
Low*	98 (28)	-		
Middle	148 (43)	0.5	0.3, 0.9	
High	99 (29)	0.4	0.2, 0.9	
Disease characteristics				
Duration of complaints				0.77
0-6 weeks*	139 (40)	-		
7-12 weeks	77 (22)	1.1	0.6, 2.1	
>3 months	134 (38)	0.9	0.5, 1.5	
Sick leave at baseline in preceding 2 months				<0.001
0 weeks*	254 (74)	-		
≤1 weeks	44 (13)	1.8	0.9, 3.9	
>1 weeks	46 (13)	3.3	1.6, 6.9	
Gradual onset (vs. acute)	212 (61)	1.0	0.6, 1.6	0.89
Precipitating cause				
Unexpected movement	16 (5)	1.7	0.5, 5.6	0.36
Strain/overuse: unusual activities	56 (16)	0.5	0.2, 1.1	0.08
Strain/overuse: usual activities	99 (28)	2.4	1.4, 4.1	<0.001
Injury	15 (4)	4..4	1.4, 15.3	0.01
Sport injury	22 (6)	0.5	0.1, 1.7	0.26
Unknown	133 (38)	1.4	0.9, 2.4	0.17
Shoulder complaints in the past	199 (57)	1.4	0.8, 2.3	0.19
Neck complaints in the past	165 (48)	1.5	0.9, 2.4	0.11
Dominant side involved	210 (60)	1.5	0.9, 2.4	0.15
Comorbid psychological complaints	27 (8)	5.4	2.0, 13.9	<0.001
Concomitant musculoskeletal complaints				
Neck/high back	119 (34)	1.6	1.0, 2.7	0.07
Low back pain	61 (17)	1.3	0.6, 2.4	0.50
Upper extremity	96 (27)	1.4	0.8, 2.5	0.21
Lower extremity	77 (22)	0.7	0.4, 1.3	0.30
Shoulder pain (0-10); mean (SD)	4.5 (2.3)			0.01
0-3 points*		-		
4-6 points		2.4	1.3, 4.5	
7-10 points		2.3	1.2, 4.5	
Shoulder disability (SDQ) (0-100); mean (SD)	58.4 (24.0)	1.9	0.8, 4.8	0.24

Table 1 continued

Variable	n (%)	OR	95% CI	p[a]
Physical examination				
ROM shoulder (0-18) (median, IQR)	3 (2-5)	0.7[#]	0.3, 1.6	0.74
Pain shoulder with movement (<12, ≥ 12 points)	6.3 (4.0)	2.9[#]	1.3, 6.7	0.03
ROM neck (0-4) (median, IQR)	0 (0-1.5)	1.2[#]	0.7, 2.2	0.41
Pain neck with movement (0-18) (median, IQR)	0 (0-3)	1.1	1.0, 1.3	0.01
Physical factors				
Physical workload (0-5) (median, IQR)	1.6 (1.6)	1.3	1.1, 1.5	0.01
Repetitive movements (yes/no)	279 (80)	1.0	0.6, 1.9	0.97
Sitting in static position for long duration (yes/no)	138 (40)	1.2	0.7, 1.9	0.54
Physical activity in comparison to others				0.09
more active*	130 (37)	-		
equally active	158 (45)	1.5	0.9, 2.7	
less active	60 (17)	2.2	1.1, 4.5	
Psychosocial factors				
Coping; mean (SD)				
Catastrophizing (1-6)	2.2 (0.8)	1.6	1.1, 2.2	0.01
Coping with pain (1-6)	2.9 (1.0)	2.6[#]	0.4, 17.9	0.44
Internal locus of control (1-6)	3.4 (0.9)	1.5[#]	0.4, 5.5	0.76
External locus of control (1-6)	3.1 (0.9)	1.9[#]	0.3, 14.2	0.76
4DSQ (median, IQR)				
Distress (0-32)	0 (0-2)	4.4[$]	1.3, 15.5	0.01
Depression (0-12)[b]	0 (0-0)	-		
Anxiety (0-24)[b]	0 (0-0)	-		
Somatization (0-32)	2 (0-4)	3.0[$]	1.0, 9.1	0.05
Fear-avoidance (0-24); mean (SD)	14.4 (5.0)	1.1	1.0, 1.1	0.04
Kinesiophobia (0-12); mean (SD)	3.3 (3.4)	1.7[#]	0.5, 5.3	0.07
Work-related psychosocial factors				
Quantitative job demands (4-20); mean (SD)	12.8 (2.7)	1.4[#]	0.6, 3.0	0.68
Skill discretion (4-20); mean (SD)	15.3 (2.8)	1.8[&]	0.6, 5.3	0.57
Decision authority (4-12); mean (SD)	9.4 (1.8)	0.8	0.7, 0.9	<0.001
Co-worker support (4-16); mean (SD)	12.3 (2.0)	10.1[&]	1.1, 92.4	0.07
Supervisor support (4-16); mean (SD)	11.2 (2.5)	2.0[&]	0.9, 4.2	0.20

SD= standard deviation; IQR = Inter quartile range; ROM = Range of Motion; 4DSQ = Four-dimensional symptom questionnaire. [a]Variables with a univariable p-value ≤0.20 were selected for the multivariable analysis. [b]ORs haven't been computed due to empty cells in the cross-tables. *Reference category. [#]In case of non-linear associations continuous variables were divided into categories. The table presents the Odds Ratio (OR) for the highest versus lowest category. [&]ORs were computed for lowest versus highest categories. [$]Variable was dichotomised.

Table 2 Secondary outcome measures at 6 months for patients with (n = 89) and without (n = 209) shoulder pain related sick leave during 6 months following first consultation

Sick leave	Yes	No	p
Persistent symptoms (%)	53	43	0.12
Pain (0-10)	2.5 (2.7)	1.8 (2.4)	0.02
Shoulder disability (SDQ) (0-100)	29.4 (30.4)	22.1 (28.8)	0.06
Severity of main complaint (0-10)	3.0 (3.3)	2.1 (2.8)	0.02

persistent symptoms, more pain ($p<0.05$), more shoulder disability, and higher severity of the main complaint at baseline ($p<0.05$).

Management of shoulder disorders

At first consultation most workers (n=253, 73%) received a wait and see policy, paracetamol, or NSAIDs. Furthermore at first consultation, 35 workers (10%) received an injection with corticosteroid, 41 workers (12%) were referred for physiotherapy and 17 workers (5%) received other therapies.

Prognostic model

The univariable associations with shoulder pain related sick leave during the 6 months following first consultation are presented in Table 1. Only variables which showed an univariable association ($p\leq0.20$) were selected for the backward stepwise selection. Table 3 presents the variables for the prediction model after backward stepwise analysis. In the backward analysis the blocks physical factors, work-related psychosocial factors, individual psychological factors, and physical examination did not result in a better model fit. A longer duration of sick leave prior to consultation, higher shoulder pain intensity, strain (overuse) as a result of usual activities and concomitant psychological complaints were associated with a higher risk of sick leave during 6 months. The multivariable regression coefficients were additionally shrunk (shrinkage factor=0.72) to obtain optimism corrected predictions for new workers.

Evaluation of the model

The reliability of the model was adequate, according to the Hosmer-Lemeshow statistic, with a non-statistically significant p-value of 0.13. Figure 1 shows the calibration of the prediction model. The plotted points are rather close to the 45° line, demonstrating good calibration over the whole range of the predictions. The distribution of the predicted risk ranges between 15 and 75% (Figure 2). The AUC of 0.70 (95% CI 0.64; 0.76) represents satisfactory discrimination.

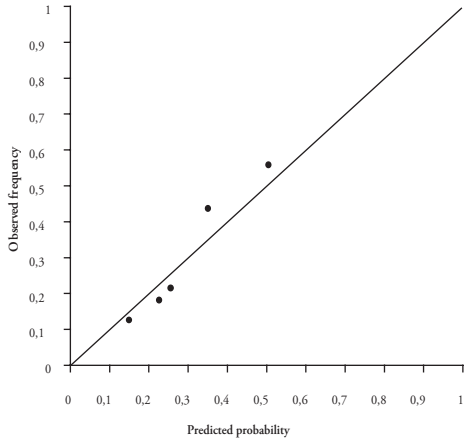
Score charts

Figure 3 shows the score chart for predicting sick leave during 6 months after first consultation. As an example, a worker with sick leave at baseline of 2 weeks (3 points), a score of 5 points on the shoulder pain scale (2 points), and shoulder complaints caused by usual activities (3 points), has a total score of 8 points corresponding to a risk of 50-60% for sick leave during 6 months after consultation.

Table 3 Multivariable model with predictors of shoulder pain related sick leave during 6 months following first consultation (n=298)

Variable	OR	95% CI
Sick leave at baseline (in preceding 2 months)		
0-weeks*		
≤1 week	1.7	0.8; 3.6
>1 months	2.2	1.0; 4.7
Shoulder pain (0-10)		
0-3 points*		
4-6 points	1.7	0.9; 3.2
7-10 points	1.9	0.9; 3.9
Strain, overuse: usual activities (yes/no)	1.9	1.1; 3.5
Concomitant psychological complaints (yes/no)	4.0	1.5; 10.8

*Reference category



Patients were grouped into quintiles according to their predicted probability of persistent symptoms according to the prediction rules. The prevalence of the endpoint within each quintiles represents the observed individual frequency.

Figure 1 Calibration plot showing the observed frequencies versus the predicted probability for sick leave in workers with shoulder pain during 6 months following first consultation.

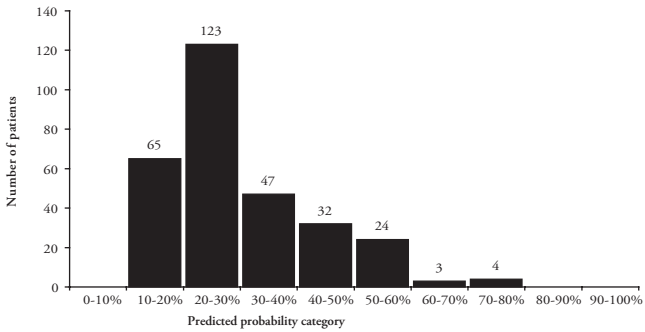


Figure 2 Distribution of predicted risk for sick leave during 6 months after first consultation (n=298)

Instruction
If a predictor is scored positively, the given weight needs to be filled in. Subsequently the scores are added to calculate the 'Total score'. From the table next to the score chart the risk (%) of sick leave for an individual patient can be determined.

Sick leave in preceding 2 months		
0-weeks	0	…
0-1 week	2	…
>1 week	3	…
Shoulder pain		
0-3 points	0	…
4-6 points	2	…
7-10 points	3	…
Strain, overuse: usual activities	3	…
Concomitant psychological complaints	6	…
		+
Total score		…

Total score	Risk
≤1	10% - 20%
2 – 3	20% - 30%
4 – 5	30% - 40%
6 – 7	40% - 50%
8	50% - 60%
9 – 10	60% - 70%
11 – 12	70% - 80%
13 – 15	80% - 90%

The predicted probability of sick leave during 6 months was determined by P=1/[1+ exp – (–1.72 + 0.53 × sick leave 0-1 week + 0.77 × sick leave >1 week + 0.50 × shoulder pain (4-6 points) + 0.65 × shoulder pain (7-10 points) + 0.68 × overuse due to usual activities + 1.38 × concomitant psychological disorders)].

Figure 3 Prognostic score chart for prediction of sick leave during 6 months following first consultation

Discussion

In this study we developed a score chart to predict shoulder pain related sick leave during 6 months following first consultation. A longer duration of sick leave prior to consultation, shoulder pain, strain (overuse) as a result of usual activities and concomitant psychological complaints were associated with a higher risk of sick leave during 6 months following consultation.

Outcome

Even though 30% of all participants reported sick leave due to shoulder pain in the 6 months following consultation, only 16% reported sick leave during at least 10 days. This seems to indicate that in our population, despite persisting pain and disability in many workers, sick leave was neither a very frequent nor long-lasting problem. An explanation could be that we did not select a high risk occupational group. Our population contained all kind of workers, irrespective of physical and psychosocial

workload. Persistent symptoms, pain, and disability were also found among those who remained at work. The differences were not very large and only significant for pain and the severity of the main complaint (Table 2).

Prognostic factors

In a systematic review[12] of the literature we found only strong evidence for age (45-54 years) as a predictor for poorer outcome. In our study no association was found for age with sick leave. It has previously been suggested that psychosocial factors such as inadequate pain cognitions and pain behaviour are likely to predict a poor outcome of painful musculoskeletal conditions.[2] Furthermore, there is evidence that the psychosocial work environment (e.g. decision authority and job satisfaction)[13] and heavy physical load (e.g. pushing and pulling, repetitive work)[13;37] may be associated with an increased risk of new episodes of shoulder pain. Our study, however, shows that these risk factors do not predict prognosis among workers who have consulted a GP for their shoulder pain. The baseline scores on psychological and psychosocial variables were generally low in our population. Significant univariable associations with sick leave during follow-up were found for several factors (distress, somatisation, catastrophising, fear avoidance, decision authority, and co-worker support), but in a multivariable model these factors had little to add to a few general and simple questions regarding the presence of psychological complaints, strain or overuse, pain intensity and sick leave at baseline. The prediction rule, consequently, contains easy to derive predictors. The prediction rule rather accurately estimates the risk of sick leave in individual workers with shoulder pain, and may help to identify workers who need additional attention.

Model fit and discrimination

The calibration plot (Figure 1) showed that the predicted probability categories were close to the ideal line. This indicates that in general the model was rather well calibrated over the complete range of predicted probabilities. The optimism corrected AUC of 0.70 implied satisfactory discrimination between shoulder patients with and without sick leave.

Generalisability

The response to the questionnaires was satisfactory in this study. Given the small differences at baseline of factors significantly associated with drop-out, we assume that the results hold for our study population. Before considering implementation of our prognostic model (i.e. score chart) in general or occupational practice the generalisability ('external validity') of the model needs to be tested in other populations of workers with shoulder pain.[38] First, the generalisability to another working population can be tested. If satisfactory, the generalisability to a community sample, or secondary care populations may be tested.

Score charts

The score charts in our study were developed to provide primary or occupational care physicians with an easy tool to predict the risk of shoulder pain related sick leave. The score charts consist of easy 'yes or no' questions, and a simple question for shoulder pain (0-10). Preferably, a physicians should use the prediction rule, programmed in a PC or PDA to calculate a risk of sick leave by answering the questions. Because we feel that not every physician has access to these equipment, we developed the score chart to enable easy implementation in everyday practice.

Reference List

1. Luime JJ, Koes BW, Hendriksen IJ, et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. Scand J Rheumatol 2004;33:73-81.
2. Van der Heijden GJMG. Shoulder disorders: a state-of-the-art review. Baillieres Best Pract Res Clin Rheumatol 1999;13:287-309.
3. Bernard B, Sauter S, Fine L, Petersen M, Hales T. Job task and psychosocial risk factors for work-related musculoskeletal disorders among newspaper employees. Scand J Work Environ Health 1994;20:417-26.
4. Skov T, Borg V, Orhede E. Psychosocial and physical risk factors for musculoskeletal disorders of the neck, shoulders, and lower back in salespeople. Occup Environ Med 1996;53:351-6.
5. Trinkoff AM, Lipscomb JA, Geiger-Brown J, Brady B. Musculoskeletal problems of the neck, shoulder, and back and functional consequences in nurses. Am J Ind Med 2002;41:170-8.
6. Bot SD, Van der Waal JM, Terwee CB, Van der Windt DAWM, Schellevis FG, Bouter LM et al. Incidence and prevalence of complaints of the neck and upper extremity in general practice. Ann Rheum Dis 2005;64:118-23.
7. Van der Linden MW, Westert GP, De Bakker DH, FG Schellevis. Complaints and disorders in the population and general practice [In Dutch: Klachten en aandoeningen in de bevolking en in de huisartsenpraktijk]. 2005 Utrecht/Bilthoven: NIVEL/RIVM.
8. Miranda H, Viikari-Juntura E, Martikainen R, et al. A Prospective study of work related factors and physical exercise as predictors of shoulder pain. Occup Environ Med 2001;58:528-34.
9. Croft P, Pope D, Silman A. The clinical course of shoulder pain: prospective cohort study in primary care. BMJ 1996;313:601-2.
10. Van der Windt DAWM, Koes BWK, Boeke AJP, et al. Shoulder disorders in general practice: prognostic indicators of outcome. Br J Gen Pract 1996;46:519-23.
11. Winters JC, Sobel JS, Groenier KH, et al. The long-term course of shoulder complaints: a prospective study in general practice. Rheumatology 1999;38:160-3.
12. Kuijpers T, Van der Windt DAWM, Van der Heijden GJGM, et al. Systematic review of prognostic cohort studies on shoulder disorders. Pain 2004;109:420-431.
13. Van der Windt DAWM, Thomas E, Pope DP, et al. Occupational risk factors for shoulder Pain: a systematic review. Occup Environ Med 2000;57:433-42.
14. Hoogendoorn WE, Bongers PM, de Vet HC, et al. High physical work load and low job satisfaction increase the risk of sickness absence due to low back pain: results of a prospective cohort study. Occup Environ Med 2002;59:323-8.
15. Ariëns GA, Bongers PM, Hoogendoorn WE, et al. High quantitative job demands and low co-worker support as risk factors for neck pain: results of a prospective cohort study. Spine 2001;26:1896-901.
16. The Dutch college of general practitioners (NHG) Practice Guideline. Accessed at http://nhg.artsennet.nl/upload/104/guidelines2/E08.htm on 1 september 2003.

17. Winters JC, Jongh AC, van der Windt DAWM, et al. Practice Guideline 'Shoulder complaints' [In Dutch: NHG-Standaard Schouderklachten]. Huisarts en wetenschap. 1999;42:222-31.

18. Karasek RA, Pieper CF, Schwart JE. Job content questionnaire and user's guide: revision 1. Los Angeles: USCLA, 1985.

19. De Jonge J, Reuvers MM, Houtman IL, et al. Linear and nonlinear relations between psychosocial job characteristics, subjective outcomes, and sickness absence: baseline results from SMASH. Study on Musculoskeletal Disorders, Absenteeism, Stress, and Health. J Occup Health Psychol 2000;5:256-68.

20. Berg SGM, Vlaeyen JWS, Ter Kuil MM, et al. Instruments for measuring chronic pain, part 2. Pain coping and cognotion list [In Dutch: Meetinstrumenten chronische pijn deel 2 Pijn Coping en Cognitie Lijst] Maastricht: Pijn Kennis Centrum, 2001.

21. Terluin B, van Rhenen W, Schaufeli W, de Haan M. The Four-Dimensional Symptom Questionnaire (4DSQ): measuring distress and other mental health problems in a working population. Work & Stress 2004;18:187-207.

22. Vendrig A, Deutz P, Vink I. Dutch version of the Fear-Avoidance Beliefs Questionnaire [In Dutch: Nederlandse vertaling en bewerking van de Fear-Avoidance Beliefs Questionnaire]. Nederlands Tijdschrift voor Pijn en Pijnbestrijding 1998;18:11-14.

23. Vlaeyen JWS, Seelen HAM, Peters M, Jong Pd, Aretz E, Beisiegel E et al. Fear of movement/(re)injury and muscular reactivity in chronic low back pain patients: an experimental investigation. Pain 1999;82:297-304.

24. Neer CS. Impingement lesions. Clin Orthop 1983;70-77.

25. Westerberg CE, Solem-Bertoft E, Lundh I. The reliability of three active motor tests used in painful shoulder disorders. Presentation of a method of general applicability for the analysis of reliability in the presence of pain. Scand J Rehabil Med 1996;28:63-70.

26. Solem-Bertoft E, Lundh I, Westerberg CE. Pain is a major determinant of impaired performance in standardized active motor tests. A study in patients with fracture of the proximal humerus. Scand J Rehabil Med 1996;28:71-78.

27. Bergman GJD, Winters JC, Groenier KH, Pool JJ, Meyboom-de Jong B, Postema K et al. Manipulative therapy in addition to usual medical care for patients with shoulder dysfunction and pain: a randomized, controlled trial. Ann Intern Med 2004;141:432-39.

28. Van der Heijden GJGM, Leffers P, Bouter LM. Shoulder disability questionnaire design and responsiveness of a functional status measure. J Clin Epidemiol 2000;53:29-38.

29. Van der Windt DAWM, Koes BWK, Deville W, et al. Effectiveness of corticosteroid injections versus physiotherapy for treatment of painful stiff shoulder in primary care: randomised trial. BMJ 1998;317:1292-96.

30. Beurskens AJ, de Vet HC, Koke AJ, Lindeman E, van der Heijden GJMG, Regtop W et al. A patient-specific approach for measuring functional status in low back pain. J Manipulative Physiol Ther 1999;22:144-48.

31. 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.

32. Efron B, Tibshirani R. An introduction to the bootstrap. Monographs on statistics and applied probability. New York: NY: Chapman and Hall; 1993.
33. Hosmer DW, Lemeshow S. Applied logistic regression. New York: john Wiley & Sons; 1989:135-75.
34. Wasson JH, Sox HC. Clinical prediction rules. Have they come of age? JAMA 1996;275:641-42.
35. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. JAMA 1997;277:488-94.
36. Wasson JH, Sox HC, Neff RK, et al. Clinical prediction rules. Applications and methodological standards. N Engl J Med 1985;313:793-99.
37. Hoozemans MJ, van der Beek AJ, Frings-Dresen MH, et al. Pushing and Pulling in Association With Low Back and Shoulder Complaints. Occup Environ Med 2002;59:696-702.
38. Justice AC, Covinsky KE, Berlin JA. Assessing the generalisability of prognostic information. Ann Intern Med 1999;130:515-24.

7

Generalisability of a prediction rule for shoulder pain related sick leave

Ton Kuijpers
Yvonne Vergouwe
Geert JMG van der Heijden
Sandra DM Bot
Jos WR Twisk
Daniëlle AWM van der Windt
Lex M Bouter

ABSTRACT

Background Recently we developed a clinical prediction rule to predict shoulder pain related sick leave during a period of 6 months after patients have consulted the GP for a new episode of shoulder pain.

Objective The objective was to evaluate the generalisability of this prediction rule by testing it in two other populations of workers consulting for a new episode of shoulder pain in primary care.

Methods A large research program, the Dutch Shoulder Study (DSS), consisting of a prognostic cohort study and three randomized trials (RCTs) with 6 months follow-up, was carried out in different geographic area's in The Netherlands. The prediction rule was derived from the results of the prognostic cohort study (n=350). The outcome was shoulder pain related sick leave during 6 months following first consultation. The control groups of the trials were merged (n=128), and used to validate the prediction rule. Besides this population, the recently conducted Musculoskeletal Disorder Study (BAS) (n=224) was used to validate the prediction rule. Generalisability of the prediction rule was tested by studying calibration and discrimination in the validation cohorts.

Results The prediction rule showed reasonable calibration in both validation cohorts. The discriminative ability, with an area under the ROC curve (AUC) of 0.70 in the derivation cohort was stable in the BAS cohort (AUC 0.71). In the control groups of the three RCTs of the DSS the discriminative ability decreased to an AUC of 0.66.

Conclusions The prediction rule for shoulder pain related sick leave during 6 months following first consultation in primary care showed adequate generalisability to another population of workers with shoulder pain participating in an observational cohort study. In the control groups of the three RCTs the prediction rule performed less well. An important next step in validating this prediction rule is to study it's applicability and predictive validity in daily practice.

Introduction

Shoulder pain is common with a one-year prevalence ranging between 5% and 47%.[1-7] In occupational settings, the one year prevalence ranges between 16% and 47%.[3-5] A Finnish study[8] reported a one-year incidence of shoulder pain of 14% among forestry workers. Shoulder pain has an unfavourable outcome in many patients. About 40 to 50% of all patients who present with a new episode of shoulder pain in primary care report persistent pain and disability after 6 to 12 months.[9-11]

Early identification (risk stratification) of patients with a high risk of shoulder pain related sick leave might enable timely intervention and prevent sick leave and the concomitant high costs these patients generate (Chapter 3). We developed a clinical prediction rule consisting of 4 easily measurable prognostic factors: sick leave at baseline in the preceding 2 months (0/0-1/≥1 weeks); shoulder pain (0-10); strain/overuse due to usual activities as a precipitating cause (yes/no); and concomitant psychological complaints (yes/no). This rule predicts shoulder pain related sick leave during 6 months after first consultation for patients with a new episode of shoulder pain in primary care. The performance (that is, calibration and discrimination) of the prediction rule was evaluated in the development study.[12] Calibration refers to what extent the observed frequencies agree with the predicted probabilities of sick leave. Discrimination refers to the ability to distinguish between a patient with a high risk of sick leave and a patient who will not have to stay off work because of shoulder pain.

Before considering implementation of the prediction rule in clinical practice the generalisability needs to be tested.[13-15] Generalisability refers to the performance in patients drawn from a different but comparable population.[13] Our objective was to evaluate the performance of our clinical prediction rule for shoulder pain related sick leave in two different cohorts of patients with shoulder pain in a primary care setting.

Methods

In this study we evaluated the generalisability of the derived prediction rule from the Dutch Shoulder Study (DSS) in a subgroup of other patients

from this cohort and among participants of another prospective cohort study in general practice; the Musculoskeletal Disorder Study (BAS).

Dutch Shoulder Study

The Dutch Shoulder Study (DSS) is a comprehensive cohort study, carried out between January 2000 and May 2005. The DSS consists of one prognostic cohort study and three randomised controlled trials, which were carried out alongside each other. Between January 2001 and June 2003, 103 general practitioners (GP) recruited patients at first consultation for a new episode of shoulder complaints in three geographic areas in the Netherlands (Amsterdam, Groningen and Maastricht). All patients in the DSS had to meet the same general inclusion criteria, and specific additional inclusion criteria if eligible for a trial (Table 1). For the prognostic cohort study no additional inclusion criteria were specified. Data from the prognostic cohort study were used to derive the prediction rule. Data from the control groups of the three trials were used to study the generalisability of the rule. For the current study only patients reporting paid work were used.

The Groningen Manipulation Study (GMO)[16;17] evaluates the effectiveness of manipulative therapy for the shoulder girdle in addition to usual care. In two other trials a Graded Exercise Therapy (GET)[18] and an Education and Activation Program (EAP)[19], respectively, were studied. Patients in the control groups of the trials received usual care, similar to the patients in the cohort study.

Baseline and follow-up assessments for all patients in the DSS were identical. The outcome was measured by postal questionnaires at 6 weeks, 3 and 6 months.

Musculoskeletal Disorder Study

The Musculoskeletal Disorder Study (BAS) is a large observational cohort study conducted in 61 general practices (97 GPs).[20;21] GPs recruited patients who consulted for a new episode of musculoskeletal pain. For our generalisability study we selected patients who consulted for shoulder pain,

and who had paid work at baseline. Selection criteria were comparable with the DSS. Follow-up questionnaires were sent after 3, 6, and 12 months.

Table 1 Selection criteria for the Dutch Shoulder Study (DSS) and Musculoskeletal Disorder Study (BAS)

General inclusion criteria DSS and BAS
Patients 18 years of age or older
Paid work
Not consulted GP or received any form of treatment for the afflicted shoulder in the preceding 3 months
Sufficient knowledge of the Dutch language
Specific inclusion criteria trials DSS
Groningen Manipulation Study (GMO)
Dysfunction of the cervicothoracic spine and adjacent ribs with accompanying pain or restricted movement
Graded Exercise Therapy Study (GET)
Duration of complaints >3 months
Education and Activation Program (EAP)
Duration of complaints <3 months
Exclusion criteria DSS and BAS
Severe physical or psychological conditions (i.e. fractures or location in the shoulder region; rheumatic disease; neoplasm; neurological or vascular disorders; dementia)

Prediction rule

The rule predicted shoulder pain related sick leave (yes ≥1 day, no=0 days) during 6 months after first consultation, and was developed using information from the 350 patients of the derivation cohort who reported paid work at baseline. Sociodemographic variables, disease characteristics (i.e. pain intensity, disability, duration of complaints, sick leave in the 2 months prior to consultation, onset, comorbidity), physical workload, work-related psychosocial factors, psychological factors and results of a physical examination were documented. The questionnaire also included a general single-item question regarding the presence (yes/no) of any psychological problems (e.g. distress, depression, anxiety). These factors were used to compose a prognostic model and derive the prediction rule. We tested the internal validity with bootstrapping techniques and corrected the prediction rule for overoptimism.[13] The calibration of the prediction rule was adequate. The discriminative ability was satisfactory with an area under the receiver operating characteristic (ROC) curve of

Instruction
If a predictor is scored positively, the given weight needs to be filled in. Subsequently the scores are added to calculate the 'Total score'. From the table next to the score chart the risk (%) of sick leave for an individual patient can be determined.

Sick leave in preceding 2 months		
0-weeks	0	…
0-1 week	2	…
>1 week	3	…
Shoulder pain		
0-3 points	0	…
4-6 points	2	…
7-10 points	3	…
Strain, overuse: usual activities	3	…
Concomitant psychological complaints	6	…
		+
Total score		…

Total score	Risk
≤1	10% - 20%
2 – 3	20% - 30%
4 – 5	30% - 40%
6 – 7	40% - 50%
8	50% - 60%
9 – 10	60% - 70%
11 – 12	70% - 80%
13 – 15	80% - 90%

The predicted probability of sick leave during 6 months was determined by P=1/[1+ exp – (–1.72 + 0.53 × sick leave 0-1 week + 0.77 × sick leave >1 week + 0.50 × shoulder pain (4-6 points) + 0.65 × shoulder pain (7-10 points) + 0.68 × overuse due to usual activities + 1.38 × concomitant psychological disorders)].

Figure 1 Prognostic score chart for prediction of sick leave during 6 months following first consultation

0.70 (95% CI 0.64+ 0.76). In Figure 1 the prediction rule is presented as a score chart. The development of this score chart and prediction rule has been described in detail elsewhere.[12]

Analysis

The performance of the prediction rule was tested in the validation cohorts by evaluating its calibration and discrimination. Calibration was assessed by plotting the predicted probabilities of sick leave according to the prediction rule, against the observed frequencies. For this, patients were grouped into quintiles according to their predicted probability of sick leave. The prevalence of the endpoint within each quintile equals the observed frequency.

The area under the ROC curve was used to assess the discriminative ability of the prediction rule. An area under the curve (AUC) of 0.5 indicates no discrimination above chance, whereas an AUC of 1.0 indicates perfect discrimination. Since the discriminative ability of a rule is related to the

homogeneity of the sample in which the rule is applied, we also estimated the maximum attainable AUC in the validation cohorts. Using the predicted risks of the patients in the validation cohorts, outcomes were generated with Monte Carlo Simulation.[22;23] The simulation mimics the situation that the model is perfectly calibrated. The AUC that is subsequently estimated for the predicted risks and generated outcomes is considered the maximum attainable AUC for the validation sample.

Furthermore, to gain insight into the performance of our prediction rule, we estimated the multivariable logistic regression coefficients for each of the predictors of our prediction rule in the validation cohorts. This analysis shows which of the different elements of the rule are the strongest predictors of sick leave in the validation cohorts.

Results

Study population

Table 2 presents the baseline characteristics of the derivation cohort and validation cohorts. Patients in the DSS control groups clearly showed a longer duration of complaints at baseline (>3 months: 51% versus 38%), and reported 10% more concomitant low back pain in comparison with the derivation cohort. Patients in the BAS were less often male (42% versus 55%), reported more often strain/overuse due to usual activities as a precipitating cause (46% versus 28%), and more often concomitant musculoskeletal complaints of the neck/high back (58% versus 34%) and low back (34% versus 17%). Patients in the derivation and validation cohorts reported similar percentages of sick leave during 6 months following first consultation of sick leave (DSS derivation cohort: 30%, DSS control groups: 34%, and BAS: 32%).

Performance

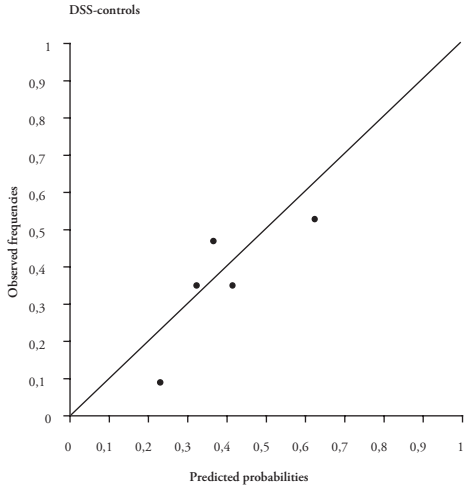
Figure 2 shows the calibration of the predictions. For the DSS control groups the predicted risks of sick leave are generally too high. Nevertheless the mean predicted probability of 0.37 was only slightly higher than the overall observed sick leave prevalence (0.32). For the BAS cohort most of

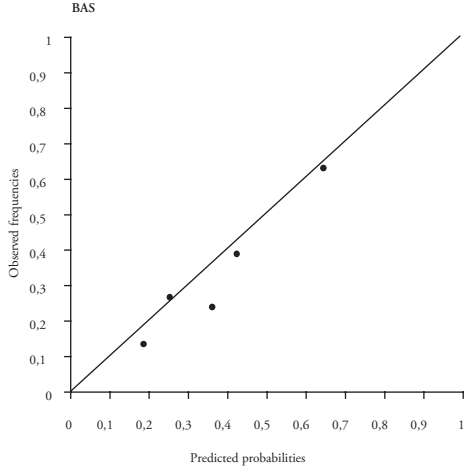
the plotted points were rather close to the 45° line, although 3 prediction categories slightly overestimated the observed probabilities. Again, the mean predicted probability (0.39) there was only slightly higher than the overall observed sick leave prevalence (0.34). The discriminative ability (AUCs) of the prediction rule was 0.66 (95% CI 0.56; 0.77) for the DSS controls groups, and 0.71 (95% CI 0.63; 0.80) for the BAS cohort. The results of the Monte Carlo Simulation show that these estimates are close to the maximum attainable AUC: 0.66 for the DSS controls, and 0.70 for

Table 2 Description of baseline characteristics in workers with shoulder pain in the derivation and validation cohorts

	Derivation	Validation	
	DSS n=350	DSS-controls n=128	BAS n=224
Demographic			
Age (years); mean (SD)	45 (11)	46 (9)	44 (10)
Gender: male; n (%)	193 (55)	63 (50)	95 (42)
Disease characteristics			
Duration of complaints >3 months; n (%)	134 (38)	65 (51)	90 (40)
Sick leave at baseline in preceding 2 months*; n (%)			
0 weeks	254 (74)	98 (78)	149 (67)
≤1 weeks	44 (13)	8 (6)	30 (13)
>1 weeks	46 (13)	20 (16)	45 (20)
Precipitating cause			
Strain/overuse: usual activities*; n (%)	99 (28)	40 (31)	104 (46)
Shoulder complaints in the past; n (%)	199 (57)	79 (62)	114 (50)
Concomitant psychological complaints*; n (%)	27 (8)	10 (8)	37 (17)
Concomitant musculoskeletal complaints; n (%)			
Neck/high back	119 (34)	51 (40)	128 (58)
Low back pain	61 (17)	34 (27)	77 (34)
Shoulder pain (0-10)*; mean (SD)	4.5 (2.3)	5.2 (2.3)	4.9 (2.2)
Work-related psychosocial factors$			
Quantitative job demands (4-20); mean (SD)	12.8 (2.7)	12.6 (2.6)	13.1 (3.2)
Decision authority (3-12); mean (SD)	9.4 (1.8)	9.6 (1.9)	9.1 (2.1)
Co-worker support (4-16); mean (SD)	12.3 (2.0)	12.3 (1.9)	12.3 (1.7)

DSS= Dutch Shoulder Study; BAS = Musculoskeletal Disorder Study; SD= standard deviation; IQR = Inter quartile range; *variables which are in the prediction rule predicting shoulder pain related sick leave; $a score between 10 and 15 points reflects fair quantitative job demands; a score > 9 points reflects high decision authority; a score >12 points reflects high co-worker support.





Patients were grouped into quintiles according to their predicted probability of shoulder pain related sick leave according to the prediction rules. The prevalence of the endpoint within each quintiles represents the observed individual frequency.

Figure 2 Calibration plots showing the observed frequencies versus the predicted probabilities for shoulder pain related sick leave during 6 months following first consultation in primary care, for the DSS-controls (n=103) and the BAS cohort (n=176)

the BAS population.

Table 3 shows the multivariable regression coefficients when the rule is applied to the validation cohorts. Shoulder pain was a strong predictor of sick leave in the DSS control groups and in the BAS cohort. In the BAS cohort sick leave in the preceding two months at baseline also showed a strong relation with sick leave during follow-up. The category '0-1 weeks' showed a remarkable negative association with outcome. A similar opposite association in both validation cohorts was seen for the category '4-6 points shoulder pain' in the BAS cohort.

Discussion

The performance of the prediction rule of shoulder pain related sick leave in the DSS control groups showed an unstable calibration and a slightly decreased discriminative ability (AUC of 0.66, compared to 0.70 in the derivation cohort). The prediction rule calibrated better in the BAS population and showed a stable discriminative ability (AUC 0.70).

The unstable calibration in the DSS control groups may partly be a result of the small numbers (n=103) in this validation cohort. Other studies have also demonstrated that calibration of prediction rules can be unstable when applied in a small population.[24]

Table 3 shows large differences between the regression coefficients in the derivation cohort and the DSS-controls. Shoulder pain was the only predictor in the DSS-controls with a large and significant regression coefficient. The decreased AUC of 0.66 was confirmed by an equal maximum attainable AUC. This decreased discriminative ability may be a result of differences in baseline characteristics between the derivation and this validation cohort. Another explanation could be that these patients origin form clinical trials, which may imply different patients characteristics regarding factors we have not documented, such as treatment preferences.

In the BAS cohort shoulder pain and sick leave at baseline in the preceding 2 months showed substantial and significant regression coefficients (Table 3). This, combined with higher regression coefficients for strain/overuse due to usual activities and concomitant psychological complaints in the

Table 3 Multivariable regression coefficients for shoulder pain related sick leave during 6 months after first consultation

	Derivation		Validation			
	DSS (n=286)		DSS-controls (n=103)		BAS (n=176)	
Predictor	β	95% CI	β	95% CI	β	95% CI
Sick leave at baseline (in preceding 2 months)						
0-weeks*						
0-1 week	0.53	-0.2; 1.3	-0.34	-3.2; 2.6	1.28	0.3; 2.2
>1 months	0.77	0.0; 1.5	0.14	-2.9; 3.1	2.66	1.5; 3.7
Shoulder pain (0-10)						
0-3 points*						
4-6 points	0.50	-0.1; 1.2	1.50	0.2; 2.8	-0.12	-1.1; 0.8
7-10 points	0.65	-0.1; 1.4	2.22	1.0; 3.5	1.13	0.1; 2.1
Strain, overuse: usual activities (yes/no)	0.68	0.1; 1.3	0.11	- 0.9; 1.1	0.38	-0.4; 1.2
Concomitant psychological complaints (yes/no)	1.38	0.4; 2.4	0.33	-1.4; 2.1	0.73	-0.3; 1.7

β= regression coefficient. The β's are derived from a multiple logistic regression analysis and shrinked using bootstrapping techniques. The β's for the validation cohorts are derived from a multiple logistic regression analysis carried out using the predictors obtained from the derivation cohort. *Reference category. CI=Confidence Interval

BAS cohort compared to the DSS-controls, may have resulted in a better performance of the prediction rule in the BAS cohort. The substantial baseline differences between the derivation cohort and the BAS, regarding factors which were not included in the prediction rule (gender, strain or overuse due to unusual activities, and concomitant musculoskeletal pain), did not seem to alter the performance. The maximum attainable AUC in the BAS cohort of 0.70 strengthens our findings of adequate discriminative ability in this cohort.

We developed a prediction rule to predict shoulder pain related sick leave during 6 months after first consultation. The elements of the prediction rule were derived from a questionnaire, filled out by the patient. If the prediction rule is used in daily practice, it is the physician who will ask the questions and calculate the risk by using a score chart. Or in a more sophisticated way, enters the responses into a personal computer (PC) or personal digital assistant (PDA), which calculates the risk of sick leave over the next six months. So, future research should also evaluate the methodologic transportability of the prediction rule (i.e. performance when data are collected by using alternative methods[13]) in a new sample of workers. And perhaps most importantly, the clinical usefulness of the instrument should be established: can the prediction rule be helpful in making decisions in the management of patients with shoulder pain, for example, whether or not to consider additional diagnostic testing, start a certain treatment or refer the patient to secondary care.[15]

In conclusion, the prediction rule for shoulder pain related sick leave during 6 months after first consultation showed disappointing generalisability in the DSS-controls and adequate generalisability in the BAS.

Reference List

1. Luime JJ, Koes BW, Hendriksen IJ, et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. Scand J Rheumatol 2004;33:73-81.
2. Van der Heijden GJMG. Shoulder disorders: a state-of-the-art review. Baillieres Best Pract Res Clin Rheumatol 1999;13:287-309.
3. Bernard B, Sauter S, Fine L, Petersen M, Hales T. Job task and psychosocial risk factors for work-related musculoskeletal disorders among newspaper employees. Scand J Work Environ Health 1994;20:417-26.
4. Skov T, Borg V, Orhede E. Psychosocial and physical risk factors for musculoskeletal disorders of the neck, shoulders, and lower back in salespeople. Occup Environ Med 1996;53:351-6.
5. Trinkoff AM, Lipscomb JA, Geiger-Brown J, Brady B. Musculoskeletal problems of the neck, shoulder, and back and functional consequences in nurses. Am J Ind Med 2002;41:170-8.
6. Bot SD, Van der Waal JM, Terwee CB, Van der Windt DAWM, Schellevis FG, Bouter LM et al. Incidence and prevalence of complaints of the neck and upper extremity in general practice. Ann Rheum Dis 2005;64:118-23.
7. Van der Linden MW, Westert GP, De Bakker DH, FG Schellevis. Complaints and disorders in the population and general practice [In Dutch: Klachten en aandoeningen in de bevolking en in de huisartsenpraktijk]. 2005 Utrecht/Bilthoven: NIVEL/RIVM.
8. Miranda H, Viikari-Juntura E, Martikainen R, et al. A Prospective study of work related factors and physical exercise as predictors of shoulder pain. Occup Environ Med 2001;58:528-34.
9. Croft P, Pope D, Silman A. The clinical course of shoulder pain: prospective cohort study in primary care. BMJ 1996;313:601-2.
10. Van der Windt DAWM, Koes BWK, Boeke AJP, et al. Shoulder disorders in general practice: prognostic indicators of outcome. Br J Gen Pract 1996;46:519-23.
11. Winters JC, Sobel JS, Groenier KH, et al. The long-term course of shoulder complaints: a prospective study in general practice. Rheumatology 1999;38:160-3.
12. Kuijpers T, Van der Windt DAWM, Van der Heijden GJMG, Twisk J, Vergouwe Y, Bouter LM. A prediction rule for shoulder pain related sick leave. Submitted.
13. Justice AC, Covinsky KE, Berlin JA. Assessing the generalisability of prognostic information. Ann Intern Med 1999;130:515-24.
14. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000;19:453-73.
15. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Validity of prognostic models: when is a model clinically useful? Semin Urol Oncol 2002;20:96-107.
16. Bergman GJ, Winters JC, Van der Heijden GJMG, Postema K, Meyboom-de Jong B. Groningen Manipulation Study. The effect of manipulation of the structures of the shoulder girdle as additional treatment for symptom relief and for prevention of chronicity or recurrence of shoulder symptoms. Design of a randomized controlled trial within a comprehensive prognostic cohort study. J Manipulative Physiol Ther 2002;25:543-9.

17. Bergman GJ, Winters JC, Groenier KH, Pool JJ, Meyboom-de Jong B, Postema K et al. Manipulative therapy in addition to usual medical care for patients with shoulder dysfunction and pain: a randomized, controlled trial. Ann Intern Med 2004;141:432-9.
18. Geraets JJ, Goossens ME, De Bruijn CP, Koke AJ, De Bie RA, Pelt RA et al. A behavioural treatment for chronic shoulder complaints: concepts, development, and study design. Aust J Physiother 2004;50:33-8.
19. De Bruijn CP, De Bie R, Geraets JJ, Goossens ME, Koke AL, Van den Heuvel HW et al. Evaluation of an education and activation programme to prevent chronic shoulder complaints: design of an RCT [ISRCTN71777817]. BMC Fam Pract 2005;6:7.
20. Van der Waal JM, Bot SD, Terwee CB, Van der Windt DAWM, Bouter LM, Dekker J. Determinants of the clinical course of musculoskeletal complaints in general practice: design of a cohort study. BMC Musculoskelet Disord 2003;4:3.
21. Bot SD, Van der Waal JM, Terwee CB, Van der Windt DAWM, Scholten RJPM, Bouter LM, et al. Predictors of outcome in neck and shoulder complaints: a cohort study in general practice. Spine (in press) 2005.
22. Gilks WR RSSD. Markov Chain Monte Carlo in practice. New York: Chapman & Hall, 1996.
23. Vergouwe Y, Steyerberg EW, Eijkemans M, Habbema J. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. J Clin Epidemiol, in press. 2005
24. Vergouwe Y, Steyerberg EW, De Wit R, Roberts JT, Keizer HJ, Collette L et al. External validity of a prediction rule for residual mass histology in testicular cancer: an evaluation for good prognosis patients. Br J Cancer 2003;88:843-7.

8

Are psychological factors associated with persistent pain and disability in patients with shoulder pain or low back pain?

Daniëlle AWM van der Windt
Ton Kuijpers
P Jellema
Geert JMG van der Heijden
Lex M Bouter

ABSTRACT

Background Psychological factors are assumed to be of prognostic importance and to predict the transition from acute to persistent pain. The influence of psychological factors has mainly been studied in low back pain, but may be equally important in other types of pain.

Objective To study the influence of psychological factors on the risk of persistent symptoms and disability at three months after consultation in patients with shoulder pain (SP) or low back pain (LBP) in general practice.

Methods Patients presenting in general practice with a new episode of SP or (sub)acute LBP were enrolled in a prospective cohort study. In both patient groups psychological factors (catastrophizing thoughts, distress, somatization, and fear-avoidance beliefs) were measured at baseline. Primary outcome measures after three months were 1) persistent symptoms, and 2) less than 30% reduction in functional disability. Logistic regression analyses were used to study the association between baseline scores on psychological factors and outcome. Interaction with symptom duration at baseline was studied for each of the psychological factors.

Results A total of 587 patients with SP and 171 patients with LBP were enrolled in the study. Drop-out rate at three months was 12% in patients with SP and 4% in patients with LBP. In patients with SP most associations of psychological factors with outcome were weak and not statistically significant. In patients with (sub)acute LBP catastrophizing thoughts and somatization were strongly and significantly associated with a higher risk of persistent symptoms and disability at follow-up.

Conclusions Psychological factors, in particular somatization and catastrophizing thoughts, are more strongly associated with persistent pain and disability in patients who consult their general practitioner for LBP than in patients with SP. This seems to indicate that different mechanisms may explain the progression to persistent symptoms among patients with different types of pain in primary care. Additional research needs to confirm these findings and further explore the role of psychological factors in the development of chronic pain problems.

Introduction

The associations between pain, disability and psychological factors have been widely studied, but mainly in cross-sectional research[1-3] or in patients with chronic pain syndromes.[4;5] Research among patients with more acute pain, preferably in a primary care setting, is needed to establish whether psychological factors also influence the progression from acute symptoms to chronic pain and disability. If these associations can be confirmed, an early identification and subsequent modification of these factors may prevent chronic disability.[6-8] Although relatively few studies have been carried out in primary care settings, a systematic review of the influence of psychosocial factors in low back pain concluded that distress and somatization are implicated in the transition from acute to chronic low back pain. Fear-avoidance and catastrophizing were mentioned as factors that deserve further investigation.[9]

So far, research on the transition from acute to chronic pain has mainly been carried out in patients with low back pain, but similar mechanisms are assumed to be of importance in other musculoskeletal pain problems, such as shoulder pain.[10] As yet, the influence of psychological factors has rarely been addressed in patients with shoulder pain. Prospective cohort studies in primary care have mainly studied the influence of disease characteristics, demonstrating that a high pain intensity at baseline is the strongest predictor of a poor outcome.[11]

The objective of this cohort study was to investigate the associations of psychological factors with the risk of persistent symptoms and disability at three months after consulting a general practitioner for a new episode of shoulder pain or low back pain. We specifically studied the influence of catastrophizing thoughts, fear-avoidance beliefs, distress, and somatization as these are assumed to be of potential importance in the development of chronic pain problems[9], and may be susceptible to intervention. We compared the influence of these factors in low back pain and shoulder pain to study whether there is similarity between different types of musculoskeletal pain regarding the mechanisms that may explain the progression to persistent pain and disability.

Methods

Study population

Shoulder pain – A total of 103 general practitioners (GP) participated in the study.[12] Patients who consulted their GP for a new episode of shoulder pain (SP) between January 2001 and June 2003 were eligible for participation. Patients were selected if they were 18 years or older, and had not consulted their GP for the afflicted shoulder in the preceding 3 months. Sufficient knowledge of the Dutch language was required to complete written questionnaires. Exclusion criteria included the presence of fractures or dislocation in the shoulder region, dementia, or systemic disease that may explain the shoulder symptoms (e.g. rheumatic disease, neoplasms, neurological or vascular disorders).

Low back pain – In the same time period participants with low back pain (LBP) were recruited by 32 GPs (21 practices) who participated in a cluster-randomised controlled trial on the treatment of low back pain in primary care.[13] Patients were selected if they were between 18 and 65 years of age, and the duration of LBP was less than 12 weeks at presentation, or there was an exacerbation of persisting back pain. Participation required sufficient knowledge of the Dutch language. Exclusion criteria were: LBP caused by specific pathological conditions (metastasis, osteoporosis, rheumatoid arthritis, or fracture), current treatment by another healthcare professional than the GP, and pregnancy. All patients receiving usual care (as their GPs had been randomised to the control group) were selected for the current study. Data collection for both cohorts was approved by the Medical Ethics Committee of the VU University Medical Center, Amsterdam, The Netherlands.

Management

All patients received usual care by their GP. For both LBP and SP the Dutch College of General Practitioners have issued national guidelines. The practice guidelines for SP recommend to provide information on the prognosis of shoulder pain, advice regarding provoking activities, and stepwise treatment consisting of analgesics, Non-Steroidal Anti-

Inflammatory Drugs (NSAIDs), corticosteroid injection or referral for physiotherapy.[14;15] Decisions regarding the content of treatment were made by the GP, dependent on the duration and severity of pain and disability. The practice guidelines for LBP advise a wait-and-see policy with pain medication and gradual uptake of activities for acute LBP (duration <6 weeks).[16] For sub-acute LBP (6 to 12 weeks duration) the guideline recommends referral for exercise therapy, physiotherapy, or manual therapy in case of persistent disability.

Psychological factors

Within a few days after recruitment by the GP all participants were visited at home by a research assistant. During the visit the participants completed a baseline questionnaire, that included questions on psychological factors. In patients with SP catastrophizing was measured with the 6-item Catastrophizing subscale of the Pain Coping and Cognition List (PCCL, 1-6 points).[17] In patients with LBP the 6-item subscale Catastrophizing of the Coping Strategies Questionnaire (CSQ, 0-36 points) was used.[18] In both populations fear-avoidance beliefs were assessed using the 4-item physical activity subscale of the Fear Avoidance Beliefs Questionnaire (FABQ; 0-24).[19] Distress as well as somatization were measured in both patient groups with two 16-item subscales of the Four-Dimensional Symptom Questionnaire (4DSQ).[20] Scores on all scales were standardized to scores between 0 and 100 to facilitate comparison. When scores were missing at least 75% of the items on subscales had to be present to calculate a total score. Otherwise, complete case analyses were performed.

Outcome measures

Outcomes were measured by postal questionnaires at 3 months after consultation. The primary outcome measures were perceived recovery and disease-specific functional disability. In both populations perceived recovery was measured on an 8-point Likert scale.[21] Patients who did not report full recovery or very much improvement were denoted as having "persistent symptoms".

For patients with SP functional disability was measured with the 16-item shoulder disability questionnaire (SDQ).[22] In patients with LBP functional disability was assessed by the 24-item Roland Morris Disability Questionnaire (RDQ).[23] Both SDQ and RDQ scores were standardized to scores ranging between 0 (no disability) and 100 (severe disability). A reduction of 30% in baseline disability score has been suggested to be a minimal clinically important change for the RDQ.[24] Research in patients with SP has shown that a reduction of[18;19] points can be considered to be a clinically important change.[25] Given a mean baseline scores of 60-75 points[12;26], this agrees with a change of 30% from baseline. Those reporting less than 30% improvement on either the SDQ or RDQ were denoted as having "persistent disability".

Other potential predictors

Psychological factors were studied as independent predictors of outcome after three months. Age, gender, pain intensity at baseline (0-10 points rating scale), duration of symptoms, and the presence of additional musculoskeletal complaints were considered as potential prognostic co-variates in both SP and LBP. Furthermore, in patients with SP a gradual or acute onset of pain, and repetitive movements (on at least two days a week) were considered as potential co-variates, whereas in patients with LBP the presence of radiating symptoms was included in the analyses. These factors have previously been shown to predict the outcome of SP[12;26-29] or LBP[30-33] in primary care populations. Finally, treatment variables were considered as potential co-variates in the association between psychological factors and outcome: corticosteroid injection in patients with SP, and referral for physiotherapy, exercise therapy or manual therapy in patients with either SP or LBP.

Theoretical model

The interactions between pain stimuli, distress, beliefs, and pain behaviour have been described in biopsychosocial models of pain.[34;35] These models provide a general explanation for the influence of psychological factors in

the transition of acute pain to chronic pain. According to such models, following an injury or nociceptive stimulus patients make a judgement concerning its meaning. These beliefs and attributions do not necessarily need to be rational (e.g. catastrophizing thoughts), and are often based on prior experiences or information obtained from others. Attributions can be influenced by distress or depressive symptoms, and lead to the development of inadequate strategies to deal with the pain (e.g. fear-avoidance behaviour). As fear-avoidance behaviours often occur in anticipation of pain rather than as a response to pain, pain behaviour may persist, resulting in more disability.[34;35]

The core notion of somatization is that psychological distress can be expressed as somatic symptoms, for which medical help is sought.[36] This concept provides a more general explanation of the association between psychological factors and persistent symptoms in patients with musculoskeletal pain.

Analysis

Descriptive statistics were used to describe baseline characteristics of patients with LBP and SP. Logistic regression analysis was used to investigate the association between psychological factors and the risk of persistent symptoms or persistent disability. Associations were expressed as odds ratios (OR) along with 95% confidence intervals (95% CI).

The following steps were taken to study the association between psychological factors and persistent symptoms or disability. First, univariable analyses were carried out to calculate the unadjusted association between each of the four psychological factors and outcome. The linearity of the associations was studied. Factors were categorized into tertiles if they did not show a linear association with the outcome. Next, these associations were adjusted for each of the other potential predictors. Only those factors that resulted in a change of at least 10% in the univariable regression coefficient of the psychological factor were considered as potential co-variates in the multivariable regression models. Next, a manual forward selection procedure was used to sequentially include co-variates in the

model that induced the largest change in the regression coefficient of the psychological factor. Co-variates that changed the multivariable regression coefficient of psychological factors by more than 10% were retained in the model. Separate models were built for patients with SP and LBP.

Finally, interaction was studied between each of the psychological factors and duration of symptoms at baseline. We considered it likely that psychological factors were stronger predictors of outcome in patients with a longer duration of symptoms (≥3 months) compared to those with a more recent onset of pain. In case of a significant association of the interaction term with the outcome ($p < 0.10$), stratified analyses were carried out for patients with short or long duration of symptoms ar baseline. All analyses were performed with the use of SPSS for Windows version 12.1 (SPSS Inc., Chicago, IL).

Results

Characteristics of the study populations

The study included 587 patients with SP and 171 with LBP. Table 1 presents the distribution of demographic variables, clinical characteristics, and psychological factors in both groups. The populations were similar with respect to gender, and educational level, but patients with LBP were younger (mean age 42.0 versus 51.5 years for SP) and more often reported paid work (84.2% versus 65.4%). Patients with shoulder pain more often reported a long duration of symptoms at baseline (at least 3 months in 41.3%), while the majority of patients with LBP reported a recent onset of their pain (less than two weeks in 61.4%). More patients with SP reported musculoskeletal pain elsewhere (71.0% versus 36.3% in those with LBP). Regardless of these differences the intensity of pain was similar in both groups.

Both groups scored equally high on fear-avoidance beliefs (median score 58.3 for SP and 62.5 for LBP), and on catastrophizing thoughts (21.7 versus 30.6). Baseline scores on distress and somatization were low in SP patients (median scores less than 7), and moderate in patients with LBP (median scores 25.0).

Recovery of symptoms after three months

Nearly all patients (164/171; 96%) with LBP completed the follow-up questionnaire at three months follow-up. Response was 88% (517/587) in patients with SP. Drop-outs with SP were younger and more often had higher distress scores (>6 points) at baseline.

Table 2 shows that more patients with SP reported persistent symptoms after 3 months (60.2% versus 45.7% in LBP). Median disability scores were higher in patients with SP, and a larger proportion showed less than 30% reduction of disability. Pain intensity scores at three months follow-up were slightly higher in LBP patients.

Table 1 Baseline characteristics of patients with shoulder pain (SP) or low back pain (LBP)

		SP (n = 587)	LBP (n = 171)
Demographic and work variables			
Age (years)	mean (SD)	51.5 (14.0)	42.0 (12.0)
Gender	n (% female)	295 (50.3)	81 (47.4)
Educational level	n (%)		
low	n (%)	210 (36.3)	56 (33.1)
middle	n (%)	234 (40.4)	72 (42.6)
high		135 (23.3)	41 (24.3)
Paid work		350 (59.6)	144 (84.2)
Repetitive movements (≥2 days/week)		384 (65.4)	-
Disease characteristics			
Duration current episode	n (%)		
0 to 2 weeks		83 (14.2)	105 (61.4)
3 to 6 weeks		122 (20.8)	49 (28.7)
6 to 13 weeks		139 (23.7)	17 (9.9)
≥ 13 weeks		242 (41.3)	-
Gradual onset	n (%)	362 (61.7)	-
Previous episodes	n (%)	348 (59.3)	138 (81.2)
Pain radiating below the knee	n (%)	-	73 (42.7)
Musculoskeletal pain elsewhere	n (%)	417 (71.0)	62 (36.3)
Pain intensity	mean (SD)	4.8 (2.3)	4.8 (1.9)
Functional disability (SDQ or RDQ, 0-100)	mean (SD)	59.9 (24.2)	50.8 (20.7)
Psychological factors			
Fear-avoidance beliefs (FABQ, 0-100)	median (IQR)	58.3 (45.8; 75.0)	62.5 (50; 79.2)
Catastrophising (PCCL or CSQ, 0-100)	median (IQR)	21.7 (12.5; 35.0)	30.6 (16.7; 41.7)
Distress (4DSQ, 0-100)	median (IQR)	0 (0; 6.3)	25.0 (12.5; 43.8)
Somatization (4DSQ, 0-100)	median (IQR)	6.3 (0; 15.6)	25.0(15.6; 34.4)

Abbreviations: SD: Standard Deviation, IQR: InterQuartile Range; SDQ: Shoulder Disability Questionnaire[22]; RDQ: Roland; Morris Disability Questionnaire[23]; FABQ: Fear Avoidance and Beliefs Questionnaire[19]; PCCL: Pain Coping and Cognitions List[17], used in patients with shoulder pain; CSQ: Coping Strategies Questionnaire[18], used in patients with low back pain; 4DSQ: Four-dimensional Symptoms Questionnaire[20]

Table 2 Persistent symptoms, pain and disability three months after consultation for patients with shoulder pain (SP) and low back pain (LBP)

		SP (n = 517) *	LBP (n = 164)*
Persistent symptoms	n (%)	311 (60.2%)	75 (45.7%)
Pain intensity (0-10)	median (IQR)	2 (0; 4)	3 (2; 5)
Disability score (0-100)	median (IQR)	30.8 (0; 60.0)	8.3 (0; 20.8)
Change in disability since baseline	mean (SD)	26.9 (31.6)	35.1 (24.7)
Less than 30% reduction in disability	n (%)	216 (41.8%)	28 (16.4%)

Abbreviations: SD: Standard deviation; IQR: InterQuartile Range *11% of SP patients (n = 70) and 4% of LBP patients (n = 7) did not return the follow-up questionnaire at three months or did not complete questions on main outcomes.

Association of psychological factors with outcome

Shoulder pain - Table 3 presents the association of baseline scores on psychological factors with persistent shoulder symptoms or shoulder disbility at three months follow-up. Associations were not linear in patients with SP, therefore psychological factors were categorised in tertiles. Given the fact that median baseline scores on distress and somatization were very low (0 for distress and 6.3 for somatization), scores on these two psychological factors were dichotomized (highest tertile versus lower scores).

Associations of all psychological factors with outcome after three months were weak, and not statistically significant. More catastrophizing thoughts at baseline were associated with a higher risk of a poor outcome at three months, but the association was not statistically significant after adjustment for baseline pain intensity and duration of symptoms.

Low back pain – The associations of psychological factors with outcome in patients with LBP were stronger, in particular for higher baseline scores on catastrophizing thoughts, distress, and somatization (Table 4). After adjustment for previous episodes of LBP, the associations with persisting symptoms were statistically significant for catastrophizing thoughts and somatization.

Interaction by symptom duration – As none of the participants with LBP reported chronic pain at baseline, interaction by symptom duration was only studied in patients with SP (Table 5). Significant interaction

Table 3 Association of psychological factors with persisting symptoms and disability after 3 months in patients with shoulder pain

		Persisting symptoms		<30% reduction in disability	
		OR	95% CI	OR	95% CI
Catastrophizing thoughts (vs < 15 points), n = 493 *					
Crude	15 to 30	0.99	0.65; 1.53	1.30	0.84; 2.01
	> 30	1.31	0.83; 2.05	1.65	1.05; 2.57
Adjusted †	15 to 30	1.02	0.63; 1.64	1.36	0.86; 2.15
	> 30	1.11	0.67; 1.84	1.62	1.00; 2.61
Distress (> 6 vs lower scores)‡, n = 514 *					
Crude		1.14	0.78; 1.66	1.07	0.73; 1.55
Adjusted †		0.73	0.47; 1.13	0.86	0.57; 1.30
Somatization (>12 vs lower scores)‡, n = 514 *					
Crude		1.50	1.03; 2.21	1.50	1.03; 2.17
Adjusted †		1.15	0.74; 1.79	1.37	0.90; 2.01
Fear-avoidance beliefs (vs < 50 points), n = 508 *					
Crude	50 to 75	1.16	0.78; 1.73	1.12	0.75; 1.67
	> 75	1.16	0.72; 1.89	1.11	0.68; 1.79
Adjusted †	50 to 75	1.25	0.81; 1.94	1.25	0.82; 1.91
	> 75	1.16	0.68; 2.00	1.21	0.72; 2.03

OR = Odds Ratio; CI = Confidence Interval; * = incidental missing values on psychological factors; † = adjusted for baseline pain intensity and duration of symptoms; ‡ = Scores on distress and somatization were very low; analysis was carried out using dichotomised scores

($p<0.10$) was only found for catastrophizing thoughts, and only for persistent symptoms as outcome measure. In patients with a long duration of symptoms at baseline (≥3 months) higher scores on catastrophizing thoughts were significantly associated with persistent symptoms at three months (OR highest versus lowest tertile: 4.47, 95% CI 1.74; 11.5), while in patients with a shorter symptom duration a reverse association was found (OR highest versus lowest tertile: 0.47, 95% CI 0.25; 0.88).

Treatment

At baseline most patients received a wait and see policy with pain medication. In the first three months 151 (26%) of SP patients received physiotherapy or manual therapy, and 99 (17%) reported treatment with local infiltration of a corticosteroid. Of the 171 patients with LBP 73

Table 4 Association of psychological factors with persisting symptoms and disability after 3 months in patients with low back pain (n=163)

		Persisting symptoms		<30% reduction in disability	
		OR	95% CI	OR	95% CI
Catastrophizing thoughts (vs < 20 points)					
Crude	20 to 40	1.89	0.89; 3.98	2.81	0.84; 9.40
	> 40	2.41	1.09; 5.36	4.88	1.47; 16.21
Adjusted †	20 to 40	2.00 ‡	0.93; 4.30	2.27 *	0.66; 7.79
	> 40	2.45 ‡	1.09; 5.51	3.31 *	0.93; 11.85
Distress (vs < 15 points)					
Crude	15 to 35	0.90	0.42; 1.94	1.40	0.42; 4.72
	> 35	2.02	0.94; 4.34	3.94	1.33; 11.71
Adjusted †	15 to 35	0.80 ‡	0.37; 1.76	1.34 †	0.39; 4.61
	> 30	1.86 ‡	0.86; 4.05	2.77 †	0.88; 8.66
Somatization (vs < 18 points)					
Crude	18 to 32	1.05	0.49; 2.26	0.83	0.28; 2.65
	> 32	3.24	1.42; 7.41	3.09	1.08; 8.82
Adjusted †	18 to 32	0.99 ‡	0.46; 2.16	0.74 †	0.22; 2.42
	> 32	3.11 ‡	1.34; 7.20	2.46 †	0.83; 7.30
Fear-avoidance beliefs (vs < 50 points)					
Crude	50 to 75	1.29	0.61; 2.71	0.84	0.28; 2.50
	> 75	1.56	0.71; 3.46	2.17	0.78; 6.02
Adjusted †	50 to 75	1.45 ‡	0.67; 3.14	0.83 †	0.27; 2.56
	> 75	1.58 ‡	0.70; 3.53	1.73 †	0.60; 5.04

OR = Odds Ratio; * = adjusted for baseline pain intensity; † = adjusted for baseline pain intensity and previous episodes of LBP; ‡ = adjusted for previous episodes of LBP

Table 5 Association of catastrophizing thoughts with persisting symptoms after 3 months: stratified analysis for patients with < 3 months or ≥ 3 months duration of shoulder pain at baseline

	Persisting symptoms	
	OR[a]	95% CI
Symptom duration < 3 months at baseline (n = 288)		
Catastrophizing thoughts (vs < 15 points)		
15 to 30	0.64	0.36; 1.13
> 30	0.47	0.25; 0.88
Symptom duration ≥ 3 months at baseline (n = 205)		
Catastrophizing thoughts (vs < 15 points)		
15 to 30	1.67	0.80; 3.51
> 30	4.47	1.74; 11.5

[a] adjusted for baseline pain intensity

(43%) were referred for therapy during follow-up period. These treatment variables did not significantly affect the reported associations between psychological factors and outcome at three months.

Discussion

This prospective cohort study among patients with SP or LBP in general practice showed a significant association between catastrophizing thoughts and somatization with persistent symptoms in patients with (sub)acute LBP. Despite longer duration of pain at presentation and a poorer prognosis, baseline scores on psychological factors were generally lower in patients with SP, and associations with outcome were weak and non-significant. Only in patients with a longer symptom duration at baseline more catastrophizing showed an increased risk of persistent shoulder symptoms at follow-up.

Influence of psychological factors in SP

The most important finding of our study is that psychological factors appeared to be of little importance in the prediction of persistent shoulder pain. So far, few studies have investigated the association between psychological factors and outcome in patients with SP, even though it has been proposed that psychological factors may be equally important in different types of pain.[10] One population-based study showed that distress predicted outcome in patients with shoulder pain[37], but these results could not be confirmed in our study. Psychological factors were of lesser importance than clinical characteristics, such as baseline intensity of pain, duration of symptoms, or additional musculoskeletal problems. One may hypothesize that in patients who consult the GP for shoulder pain the symptoms are more severe than in a population-based sample, perhaps resulting from more serious physical pathology, and that in these patients psychological factors may be less important.

A recently conducted prospective cohort study among patients with neck-shoulder pain in Dutch primary care showed that worrying was associated with poor outcome after three months follow-up.[27] Further research may

establish whether worrying or distress are especially important among patients with neck pain, or combined neck-shoulder pain, compared to those with shoulder pain only.

Influence of psychological factors in LBP

Our results strengthen the findings of previous research among patients with LBP, reporting significant associations between high levels on catastrophizing thoughts or somatization at baseline and the risk of persistent symptoms at follow-up.[9;30;38;39] Although frequently suggested as a potentially important predictor, associations between fear-avoidance beliefs and persistent symptoms were weak. This finding seems to coincide with the results of a prospective cohort study by Burton et al. among primary care patients with LBP.[39] Apparently, fear-avoidance beliefs do not contain unique predictive qualities independent of other prognostic factors in these primary care populations.

As we were specifically interested in the predictive value of each psychological factor, we adjusted the associations for other prognostic factors, such as pain intensity and previous episodes of LBP associations. Following adjustment most associations were weaker, and not statistically significant. One might argue that adjustment for pain intensity or a history of pain is not sensible, as the level of distress, somatization or catastrophizing is likely to depend on the duration, history or intensity of pain.[40] Longitudinal research with repeated assessment of psychological factors, symptoms, and disability in patients with acute pain is needed to address these assumptions, and further unravel the longitudinal associations between psychological factors and pain.

Comparison of the influence of psychological factors in SP and LBP

Research in which the predictive value of psychological factors is simultaneously studied in patients with different types of musculoskeletal pain is scarce. Most cohorts have used different questionnaires for measuring predictors and outcomes, hampering the comparison across different patient populations. Our study showed that the two populations

were clearly different at baseline: due to different selection criteria the mean duration of symptoms was longer in SP patients. Furthermore, SP patients reported more additional musculoskeletal pain and somewhat higher scores on baseline disability. Consequently, one might have expected higher scores on psychological factors. Yet, baseline scores were lower compared to patients with LBP, in particular for distress and somatization, and hardly influenced outcome. More patients with LBP reported having had previous episodes of LBP. It may be possible that psychological factors become more important once patients have experienced more episodes of pain. Inadequate beliefs and attributions of pain may become stronger when recurrences of pain occur.

High scores on catastrophizing thoughts predicted a poor outcome in patients with (sub)acute LBP and also in patients with longstanding SP (duration more than 3 months at consultation). Catastrophizing is an exaggerated orientation towards pain stimuli, and is considered to be an inadequate coping strategy.[41] This characteristic could be addressed by cognitive behavioural interventions. Future research might be aimed at the development and evaluation of interventions aimed at reducing catastrophizing thoughts in patients with LBP or chronic SP in primary care. The finding that catastrophing thoughts appeared to be associated with a favourable prognosis in patients with more acute SP is surprising, and difficult to explain. It is the question whether these findings can be replicated in future research among primary care patients with SP.

Methodological considerations

The study among patients with LBP was characterized by a very low dropout rate and few missings. Dropout rate in the SP cohort was limited to 11%. Although dropouts were younger and more often had higher distress scores, the absolute differences in age and distress scores were small (less than 10 years, and less than 4 points for distress). Therefore, we do not believe that these differences have strongly influenced the reported associations between psychological factors and outcome in patients with SP.

The size of the SP cohort was more than adequate to detect relevant associations with statistical significance. As the LBP cohort was smaller, it is possible that we have missed relevant associations between some of the psychological factors and outcome in patients with LBP. Power calculations for observational research are difficult to make, but a frequently used general rule is that 10 cases (events) are needed for each variable in the model.[42] With 75 patients reporting persistent LBP in our cohort we had sufficient power to establish statistically significant associations for one (categorized) predictor, with adjustment for two to three co-variates. We consider it unlikely that a possible lack of power explains the absence of significant associations for fear avoidance beliefs, as associations were rather weak.

The SP patients participated in an observational study, giving a fair representation of patients consulting their GP for a new episode of (non-traumatic) SP. Patients with LBP had been enrolled in a randomised trial, which usually generates a more selective and homogeneous study population. In this trial, however, no specific selection criteria other than duration of symptoms had been used. Furthermore, randomisation took place at the level of the practice, which means that GPs in the control group simply provided usual care to all LBP patients in their practice, and had not been trained to provide any other treatment. The participants, therefore, do represent patients with non-specific (sub)acute LBP in general practice, which increases the generalisability of our findings.

In conclusion, psychological factors, in particular somatization and catastrophizing thoughts, are more strongly associated with persistent pain and disability in patients with LBP than in those with SP. Despite a longer duration of pain at presentation and a poorer prognosis, scores on most psychological factors were low in patients with SP and associations with outcome were weak. This seems to indicate that different mechanisms may explain the progression to persistent symptoms among patients with different types of pain. Further research in primary care populations is needed to confirm these findings, and to look more closely at the role of psychological factors in the development of persistent pain and disability.

Reference list

1. Verbunt JA, Seelen HA, Vlaeyen JWS, Van der Heijden GJMG,, Knottnerus JA. Fear of injury and physical deconditioning in patients with chronic low back pain. Arch Phys Med Rehabil 2003;84:1227-32.
2. Woby SR, Watson PJ, Roach NK, Urmston M. Adjustment to chronic pain: the relative influence of fear-avoidance beliefs, catastrophising, and appraisals of control. Behav Res Ther 2004;42:761-74.
3. Badcock LJ, Lewis M, Hay EM, McCarney R, Croft PR. Chronic shoulder pain in the community: a syndrome of disability or distress? ARD 2002;61:128-31.
4. Turk DC, Okifuji A. Psychological factors in chronic pain: evolution and revolution. J Consul Clin Psychol 2002;70:678-90.
5. Crombez G, Vlaeyen JWS, Heuts PHTG, Lysens S. Fear of pain is more disabling than pain itself. Evidence on the role of pain-related fear in chronic back pain disability. Pain 1999;80:329-340.
6. Pincus T, Vlaeyen JWS, Kendall NAS, Von Korff MR, Kalauokalani DA, Reiss S. Cognitive-behavioural therapy and psychosocial factors in low back pain. Directions for the future. Spine 2002;27:E133-8.
7. Linton SJ. Early identification and intervention in the prevention of musculoskeletal pain. Am J Ind Med 2002;41:433-42.
8. Linton SJ, Boersma K. Early identification of patients at risk of developing a persistent back problem: the predictive value of the Orebro Musculoskeletal Pain Questionnaire. Clin J Pain 2003;19:80-6.
9. Pincus T, Burton AK, Vogel S, Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. Spine 2002;27: E109-120.
10. Linton SJ. An overview of psychosocial and behavioural factors in neck-and-shoulder pain. Scand J Rehabil Med Suppl 1995;32:67-77.
11. Kuijpers T, van der Windt DAWM, van der Heijden GJMG, Bouter LM. Systematic review of prognostic cohort studies on shoulder disorders. Pain 2004;109:420-431.
12. Kuijpers T, Van der Windt DAWM, Boeke AJP, Twisk JWR, Vergouwe Y, Bouter LM, Van der Heijden GJMG. Clinical prediction rules for the prognosis of shoulder disorders in general practice. Pain, in press.
13. Jellema P, van der Windt DAMW, van der Horst HE, Twisk JWR, Stalman WAB, Bouter LM. Should treatment of (sub)acute low back pain be aimed at psychosocial prognostic factors? Results of a cluster-randomised clinical trial in general practice. BMJ [in press].
14. The Dutch College of General Practitioners (NHG) Practice Guideline 'Shoulder conmplaints'. Available at: http://nhg.artsennet.nl/upload/104/guidelines2/E08.htm. 1999.
15. Winters JC, Jongh AC, van der Windt DAWM, et al. NHG-Standaard Schouderklachten. Huisarts en wetenschap 1999;42:222-231.
16. Faas A., Chavannes AW, Koes BW, Van den Hoogen JMM, Mens JMA, Smeele LJM, Romeijnders ACM, van der Laan JR. NGH-Standaard Lage-Rugpijn. Huisarts en Wetenschap 1996;39:18-31.

17. Berg SGM, Vlaeyen JWS, Kuil MM ter, Spinhoven P, Breukelen G van, Kole-Snijders AMJ. Meetinstrumenten chronische pijn deel 2. Pijn Coping and Cognitie Lijst. Maastricht: Pijn Kennis Centrum, 2001.

18. Rosenstiel AK, Keefe FJ. The use of coping strategies in chronic low back pain patients: relationship to patiënt characteristics and current adjustment. Pain 1983;17:33-44.

19. Waddell G, Somerville D, Henderson I, Newton M, Main CJ. A fear-avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain 1993;52:157-168.

20. Terluin B, van Rhenen W, Schaufeli WB, de Haan M. The four-dimensional symptom questionnaire (4DSQ): measuring distress and other mental health problems in a working population. Work & Stress 2004;18:187-207.

21. Buchbinder R, Bombarider C, Yeung M, Tugwell P. Which outcome measures should be used in rheumatoid arthritis trials? Arth Rheum 1995;38:1568-80.

22. Van der Heijden GJMG, Leffers P, Bouter LM. Shoulder disability questionnaire design and responsiveness of a functional status measure. J Clin Epidemiol 2000;53:29-38.

23. Roland M , Morris R. A study of the natural history of back pain. Part I: Development of a reliable and sensitive measure of disability in low back pain. Spine 1983;8:141-144.

24. Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1. Phys Ther 1998;78:1186-96.

25. Van der Windt DAWM, Van der Heijden GJMG, Koes BW, Winter AF de, Devillé WLJM, Bouter LM. The responsiveness of the shoulder disability questionnaire. ARD 1998;57:82-87.

26. van der Windt DA, Koes BW, Boeke AJ, Devillé W, De Jong BA, Bouter LM. Shoulder disorders in general practice: prognostic indicators of outcome. Br J Gen Pract 1996;46:519-23.

27. Bot SDM, Van der Waal JM, Terwee CB, Van der Windt DAWM, Scholten RJPM, Bouter LM, Dekker J. Predictors of outcome in neck and shoulder symptoms: a cohort study in general practice. Spine [in press].

28. Winters JC, Sobel JS, Groenier KH, Arendzen JH, Meyboom-de Jong B. The long-term course of shoulder complaints: a prospective study in general practice. Rheumatology 1999;38:160-3.

29. Croft P, Pope D, Silman A. The clinical course of shoulder pain: prospective cohort study in primary care. Primary Care Rheumatology Society Shoulder Study Group. BMJ 1996;313:601-602.

30. Klenermann L, Slade PD, Stanley M, Pennie B, Reilly JP, Atchinson LE, Troup JDG, Rose MJ. The prediction of chronicity in patients with an acute attack of low back pain in a general practice setting. Spine 1995;20:478-84.

31. Flynn T, Fritz J, Whitman J, et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. Spine 2002;27:2835-43.

32. Thomas E, Silman AJ, Croft PR, Papageourgiou AC, Jayson MIV, Macfarlane G. Predicting who develops chronic low back pain in primary care: a prospective study. BMJ 1999;318:1662-7.

33. Von Korff M, Deyo RA, Cherkin D, Barlow W. Back pain in primary care. Outcomes at 1 year. Spine 1993;18:855-62.
34. Waddell G. A new clinical model for the treatment of back pain. Spine 1987;12:632-44.
35. Vlaeyen JWS, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000;85:317-32.
36. Lipowski ZJ. Somatization: the concept and its clinical application. Am J Psychiatry 1988;145:1358-68.
37. Macfarlane GJ, Hunt IM, Silman AJ. Predictors of chronic shoulder pain: a population based prospective study. J Rheumatol 1998;25:1612-5.
38. Dionne CE, Koepsell TD, Von Korff M. et al. Predicting long-term functional limitations among back pain patients in primary care settings. J Clin Epidemiol 1997;50:31-43.
39. Burton AK, Tillotson KM, Main CJ, Hollis S. Psychosocial predictors of outcome in acute and subchronic low back trouble. Spine 1995;20:722-7.
40. Sieben JM, Vlaeyen JWS, Tuerlinckx S, Portegijs PJM. Pain-related fear in acute low back pain: the first two weeks of an episode. Eur J Pain 2002;6:229-37.
41. Sullivan MJL, Bishop S, Pivik J. The pain catastrohising scale: development and validation. Psychol Assessment 1995;7:524-32.
42. Altman D. Practical statistics for the medical research. London: Chapman & Hall, 1991.

9

General discussion

The main objectives of this thesis were to derive and validate clinical prediction rules to predict persistent shoulder symptoms at 6 weeks and 6 months after first consultation, and to predict shoulder pain related sick leave during 6 months following first consultation. Furthermore, this thesis presents a systematic review of the current literature on the prognosis of shoulder pain, an estimation of the total costs during 6 months after first consultation, and a study of the association between psychological factors and the persistence of shoulder pain.

In this chapter the main findings will be summarized and the design and results of this prognostic study will be critically discussed. Subsequently the systematic review, costs of shoulder pain, the design and conduct of the prognostic cohort study, the development and generalisability of the prediction rules and the role of psychosocial factors will be critically reviewed. At the end of this chapter final conclusions and recommendations for future research are given.

Systematic review of the literature

In our systematic review we found disappointingly little evidence for most factors which in the current literature are suggested to be of prognostic importance. There is consistent evidence that high baseline pain intensity in primary care populations and age between 45 and 54 years in occupational populations are strong predictors of a poor prognosis, while there is some evidence that a long duration of complaints and a high disability score at baseline are predictors of a poor prognosis in primary care populations. There were no studies of sufficient quality in secondary care. No evidence for the prognostic importance of psychosocial factors was found in our systematic review.

Besides methodological shortcomings there was considerable heterogeneity regarding design, study populations, prognostic factors and outcome measures. This heterogeneity impedes meta-analysis. Therefore we decided to perform a best evidence synthesis of the available evidence. Any system for defining levels of evidence is arbitrary. We chose a system that has been used in a systematic review on prognostic factors for whiplash

related disorders[1]. Systematically reviewing prognostic studies is still in development and no validated or widely used criteria list is available. Moreover, because of the small number of studies on which our conclusions were based, and the heterogeneity across studies regarding duration of follow-up, types of outcome measures, and analysis, we feel that the results of our review need to be interpreted with considerable caution. In our opinion an appropriately designed and reported prospective cohort study should fulfil the criteria of our checklist (Chapter 2). Estimates of outcome should preferably be expressed as absolute risks, and not as RRs or ORs.

Our prognostic study earned 17 points (94% of the total score) on the criteria list and was of high quality according to our a priori definition. If we add our prognostic study to the best evidence synthesis of the systematic review, in primary care populations the evidence for 'high pain intensity at baseline' as important predictor would be confirmed as being strong. The evidence for 'long duration of complaints' would change from moderate into strong. For occupational populations evidence for the prognostic value of 'sick leave at baseline' and 'high pain intensity' would change from inconclusive to moderate evidence.

Costs of shoulder pain

The present study is the first to evaluate the overall costs generated by patients presenting with shoulder pain in primary care. With a mean total costs of € 689 per patient during the 6 months after first consultation, healthcare consumption and sick leave did not seem to be very high in this primary care population. A small part (12%) of the population accounted for 74% of the total costs.

An explanation for the on average modest health care costs could be that general practitioners stick to the interventions recommended by the Dutch guidelines for shoulder disorders[2;3] (wait-and-see policy with pain medication, followed by injections), which are relatively inexpensive. Indirect costs accounted for a large proportion (47%) of the total costs. Nevertheless, the total number of days sick leave per patient was small (2.8 days) over a period of 6 months. Possibly, factors such as shoulder pain,

sleeping problems, or loss of function have caused loss of productivity in patients without sick leave from paid work. Our study does not provide information on the actual loss of productivity among those who kept on working regardless of their shoulder pain.

Similar to studies on low back pain[4], in our study a small proportion of the population (12%) caused a substantial part (74%) of the total costs. In this subgroup sick leave from paid work accounted for 61% of the total costs. In our study we were able to include a follow-up period of 26 weeks. It is possible that prolonged and recurrent pain episodes generate additional costs for more expensive care, e.g. more physiotherapy, diagnostic imaging and surgical treatment, including hospitalisation. Given the poor prognosis of shoulder pain (approximately 40% of patients report persistent symptoms after 12 to 18 months[5-7]) higher health care costs and productivity losses may be expected when follow-up times are longer. In the 6 months following first consultation, few costs were made due to referrals to other health care providers, additional diagnostic procedures, or surgery. We expect these kind of health care expenses to occur in the long-term in a small subset of the population. In our cohort patients with fractures, dislocation, or previous surgery were not included. These patients may also generate substantial costs.

Given the high incidence of shoulder pain (12/1000/year)[8-10] in general practice the total costs to society can be substantial. Future studies on the cost of illness of shoulder pain should include a longer follow-up period, and include patients with fractures, dislocation, or previous surgery to estimate the total costs of illness for patients consulting with shoulder pain in primary care.

Prognostic cohort study

Patient selection

We conducted a prospective cohort study in three geographic areas in The Netherlands. Patients who visited their general practitioner with a new episode of shoulder pain were included in the prognostic cohort study. We defined a new episode as a patient who had not consulted his/her GP or

received any form of treatment for the afflicted shoulder in the preceding 3 months.

In this large prognostic cohort study of 587 patients, the response to the follow-up questionnaires was high (between 83% and 92%). Given the low drop-out rate and only slight differences at baseline between drop-outs and responders we consider the results to be valid. Of the 587 participants, 41% reported a duration of their complaints at baseline of more than 3 months, and 62% reported that they had experienced shoulder complaints in the past. These numbers indicate that we did not enrol many patients who experienced shoulder pain for the first time in their lives, which potentially has had an influence on our prognostic models. However, patients did not seek any medical care in the preceding three months before inclusion, indicating a new episode of shoulder pain, which we feel is important for a prognostic cohort study.

Management

The ideal situation when conducting a prognostic cohort study is that patients do not receive any form of treatment. Because of the ethical problems when conducting such a study, and the expectation that not many patients would be interested to participate, this is not a feasible option. The second best option is to standardise treatment and have all participants treated in the same way. A third option is to register the management regime as good as possible.

In our study all patients received (standardised) treatment according to the 1999 version of the Dutch guidelines for shoulder disorders issued by the Dutch College of General Practitioners.[2;3] The guidelines recommend giving information on the prognosis of shoulder pain, advice regarding provoking activities, and stepwise treatment consisting of paracetamol, Non-Steroidal Anti-Inflammatory Drugs (NSAIDs), corticosteroid injection or referral for physiotherapy. The GP made the decision regarding the content of treatment based on duration and severity of pain and disability. The participating general practitioners were educated and trained to apply treatment according to this guideline.

We decided not to select treatment as a potential predictor for our prediction models. Including treatment options in a prediction rule may provide confusing information: that is a risk of persistent symptoms, provided that the patient receives a certain treatment. This means that a GP could influence the calculated risk derived from the prediction rule, by choosing a certain treatment. Treatment as a predictor in our rule would seem to imply that we have studied the effect of a certain intervention, which was no objective in our study (since our study was not a randomised controlled trial). Our prediction rules have been developed with the objective of risk stratification, and are based on information gathered through history taking and physical examinations. Subsequent decisions regarding treatment can be based on this information. Information regarding effects of treatment have possibly been influenced by confounding by indication. Patients with more severe symptoms and thus, with a higher risk of a poorer outcome were probably more likely to receive more extensive treatment.[11] Only 68 patients (12%) received an injection and 58 (10%) were referred to a physiotherapist, which is a low proportion compared to an earlier study in The Netherlands.[5] The Dutch practice guidelines on shoulder complaints, which recommend a wait-and-see policy during the first 2 to 4 weeks may have lead to a change in practice over the past 5 years. As most patients received wait and see policy or medication, we had a relatively homogeneous group regarding treatment at baseline.

Outcome assessment

Outcome was measured by postal questionnaires at 6 weeks, 3 and 6 months. In our study we focussed on persistent symptoms, and we chose as primary outcome measure patient perceived recovery, which was measured on a 8-point Likert scale[12;13] Patients who did not report full recovery or very much improvement were denoted as having persistent symptoms. We feel this outcome measure reflects daily practice, in which questions are asked by physicians or therapists like 'Has your shoulder pain recovered?', 'How is your shoulder pain?', or 'Do you still have shoulder complaints?'. At 6 weeks 70% (340/486) and at 6 months 46% (249/536) patients reported persistent symptoms, which are similar percentages to those

found in other studies carried out in primary care populations.[5-7] This may strengthen the generalisability of our findings to other primary care patients with shoulder pain. Secondary outcome measures were shoulder disability, pain, and severity of the main complaint. Our results show that patients with persistent symptoms at follow-up reported also more pain, more shoulder disability and higher severity of the main complaint than patients reporting improvement of symptoms. This strengthens the choice of our main outcome measure.

For studying shoulder pain related sick leave 350 patients who reported paid work were used, of whom 30% (89/298) reported sick leave during 6 months after first consultation, which was similar to the percentages found in the validation cohorts of our generalisability study. The duration of sick leave was limited to a few days in most workers. Our data were not suitable for studying predictors of long-lasting sick leave or frequent absenteeism as a consequence of the low incidence of these events.

Prediction rules

We developed clinical prediction rules consisting of a limited number of (easily measurable) prognostic factors to predict the risk of persistent shoulder symptoms at the short (6 weeks) and long term (6 months) and to predict shoulder pain related sick leave during 6 months following first consultation. We developed the rules in the prognostic cohort study of the Dutch Shoulder Study, and subsequently evaluated their generalisability by testing them on the merged control groups of the three trials embedded in the DSS. The prediction rule for shoulder pain related sick leave was also tested in participants of the Musculoskeletal Disorder Study (BAS).

Development

During the development of our prediction rules it was not possible to select predictors based on clinical knowledge, because there was insufficient evidence regarding their predictive value.[14] So we had to base the selection of potential predictors on univariable analyses. We selected predictors showing statistically significant associations with

the outcome using a relative high p-value of 0.20, in order not to miss any potentially relevant predictors for our multivariable analyses. In the current literature bootstrapping techniques are considered to be the most adequate for studying the internal validity of prediction models. We used these techniques to study the internal validity of the final prediction model, i.e. to adjust the estimated regression coefficients for overfitting and the model performance for overoptimism.[15;16] The performance of the prediction rules was evaluated by studying calibration and discrimination. Calibration refers to what extent the observed frequencies agree with the predicted probabilities. Discrimination refers to the ability to distinguish between a patient with or a patient without persistent symptoms or sick leave.

The calibration of the prediction rules for persistent symptoms in the derivation cohort was adequate. The discriminative ability was satisfactory for the short term (6 weeks) with an area under the curves (AUC) of 0.74 and disappointing for the long term (6 months) and 0.67 at 6 months. The calibration of the prediction rule for shoulder pain related sick leave was adequate. The discriminative ability of the prediction model was satisfactory with an AUC of 0.70.

Generalisability

Before considering implementation of the prediction rules in clinical practice their generalisability needs to be tested.[17-19] Generalisability refers to the performance of the prediction rules in patients drawn from a different but comparable population.[17]

In general, the prediction rules showed better calibration in the derivation cohort than in the validation cohorts. This is not surprising because the models were fit on the data of the derivation cohort. Since the discriminative ability of a rule is related to the homogeneity of the sample in which the rule is applied, we estimated the maximum attainable AUC. Given the predicted risks of the patients in the validation cohort, outcomes were generated with Monte Carlo Simulation.[20;21] This simulation mimics the situation that the model is perfectly calibrated. The AUC that is

subsequently estimated for the predicted risks and generated outcomes is considered the maximum attainable AUC for the validation sample.

The performance of the prediction rule for persistent shoulder symptoms at 6 weeks in the validation cohort was satisfactory, with a reasonable calibration, and discriminative ability similar to that found in the derivation cohort (AUC=0.72). The prediction rule for persistent symptoms at 6 months showed poor calibration and discrimination. The AUC decreased to 0.57 in the validation cohort, not much more than a flip of a coin (AUC=0.50).

The prediction rule for shoulder pain related sick leave was validated in the merged control groups of the trials of the DSS, and in the Musculoskeletal Disorder Study (BAS). The prediction rule showed reasonable calibration in both validation cohorts. The discriminative ability was stable in the BAS (AUC 0.71 compared to an AUC of 0.70 in the derivation cohort). In the controls of the DSS the discriminative ability decreased to an AUC of 0.66.

In the 10 day period between first consultation and baseline assessment symptoms have changed, which may have resulted in a different population at baseline assessment compared to the moment of GP consultation. This could have had an influence on the discriminative ability of our prediction rules. We do not have information that provides insight into the extent to which symptoms have changed between selection and baseline assessment, so we cannot estimate this influence on the performance of our prediction rules. This is another important reason why we want to stress the importance of validating the prediction rules in a daily practice situation, for which they have been developed.

In conclusion, we think these results are promising to test the performance of the prediction rules in a daily practice setting. The performance of the prediction rule for the long term (persistent symptoms at 6 weeks) in a daily practice situation is questionable given its poor performance in the derivation and validation cohort.

Predictors

A longer duration of symptoms at baseline, gradual onset of shoulder complaints, and more severe pain intensity were associated with a poorer prognosis at both 6 weeks and 6 months. Furthermore, concomitant psychological complaints, repetitive movements, and more severe neck pain scores at physical examination were associated with persistent symptoms at 6 weeks. A poor prognosis at 6 months was additionally predicted by concomitant back pain and more severe shoulder pain scores at physical examination. In line with our systematic review of the literature we could confirm the association between more severe pain intensity and outcome.

A longer duration of sick leave prior to consultation, more severe shoulder pain, strain (overuse) as a result of usual activities and concomitant psychological complaints were associated with a higher risk of sick leave during 6 months following first consultation. We could not confirm the strong evidence found in our review for age between 45 and 54 years as a predictor of poor outcome.

An important message of this study is that the strongest predictors of outcome were disease characteristics, results of physical examination, a single work-load factor, and a single yes or no question regarding concomitant psychological complaints.

Psychological factors

A priori we expected associations between psychological (such as distress, somatization, fear and avoidance) and work related psychosocial factors (such as decision authority, job demands) in predicting the outcome of shoulder pain in primary care. No psychological factors where selected in our prediction models, except a single yes or no question regarding concomitant psychological complaints. For the applicability of the prediction rules in primary care this is an advantage, because in a prediction rule easy to measure predictors are preferred above predictors which are measured with time taking full-length questionnaires. In chapter 8 we specifically studied the role of psychological factors, and found that in particular somatization and catastrophizing thoughts play a more

important role in the transition from acute to chronic pain in low back pain than in shoulder pain. Despite more chronic pain at presentation and a poorer prognosis, baseline scores on most psychological factors were lower in patients with shoulder pain and associations with outcome were weak and non-significant. Only in patients who already reported chronic shoulder pain at baseline catastrophizing thoughts predicted a poor outcome at follow-up.

Absolute risks

As shown in our review most prognostic studies present their results in terms of RRs or ORs. For the patient (and the physician) these measures are not easy to interpret. We feel it is an important strength of our study that with the prediction rules rule absolute risks are calculated.

Score charts

In this thesis we developed three score charts. With this charts the physician can calculate the risk of persistent shoulder symptoms at 6 weeks and 6 months after consultation and the risk of shoulder pain related sick leave during 6 months following first consultation. The sum of the scores correspond with the risk of a poor outcome. We present our prediction rules as score charts because we think this as an easy-to-use tool, which facilitates the calculation of an individual patient's risk in daily practice.

It is important to present the predicted risks with some estimate of uncertainty. It is possible to present 95% Confidence Intervals (CI) along with the calculated absolute risk derived with the prediction rule. This will provide the physician with information about the precision of the estimate. We decided to present the calculated risks in categories of at least 10%. These categories incorporates some uncertainty of the estimates, but are easy to understand and convey to the patient. As possible answers to questions of patients regarding his or her prognosis we have phrases in mind like: "You have 20-30% risk that your symptoms will persist for at least 6 weeks".

Conclusions

- In our systematic review we found little evidence for most factors which in the current literature are suggested to be of prognostic importance in patients with shoulder pain.
- The total costs (€ 689) in the 6 months following first consultation in primary care are not very high. Only a small proportion (12%) of the population generated 74% of the total costs.
- The prediction rule for the short term (6 weeks) prognosis of shoulder pain in general practice showed satisfactory performance in the derivation cohort and adequate generalisability to another but comparable population.
- Predicting long term (6 months) outcomes of shoulder pain in general practice is difficult; our prediction rule was not very successful in discriminating between patients with or without persistent symptoms.
- The prediction rule for shoulder pain related sick leave showed satisfactory performance in the derivation cohort and adequate generalisability to another primary care population with shoulder pain. It performed less well in the merged control groups of the DSS.
- Psychological factors hardly influence the risk of a poor outcome in patients with shoulder pain.

Recommendations

Future research

It is important to validate the developed prediction rules in a new set of patients and test their performance in daily practice. In other words, our prediction rules should be tested when the GP at first consultation calculates the risk of persistent symptoms at 6 weeks and optionally at 6 months, respectively the risk of shoulder pain related sick leave during the 6 months following first consultation.

Most importantly, the clinical usefulness of the rules should be established: can the prediction rules be helpful to the clinician when making decisions

in the management of patients with shoulder pain, for example, whether or not to consider additional diagnostic testing, start a certain treatment or refer the patient to secondary care.[19] For this, the prediction rules could be tested in a randomised controlled trial. The question could be studied whether a subgroup having a high risk of persistent symptoms according to the prediction rule would benefit more from a certain treatment than patients with a low risk of persistent symptoms. Such information could have an influence on daily practice. Patients with low risks of persistent symptoms could receive a wait-and-see policy, while patient with high risks of persistent symptoms would receive more extensive treatment, i.e. injection, physical therapy, or referral to secondary care.

Daily practice

The prediction rule can be helpful in daily practice for GPs or other physicians to obtain an estimation of a patients prognosis of shoulder pain or the risk of shoulder pain related sick leave. Furthermore, it can provide patients with an adequate answer to questions regarding their prognosis as 'How long will my shoulder pain persist?' or 'When will I be recovered?'.

The prediction rule for sick leave is not developed from a patients perspective, but to provide physicians insight in the risk of a patient reporting shoulder pain related sick leave in the 6 months following first consultation for shoulder pain.

Futuristic recommendations

In the Appendix we present a digital score chart, which combines the three prediction rules of this thesis. In the future following history taking and physical examination, a GP enters the responses to the questions of the prediction rules (Appendix) into a personal computer (PC) or personal digital assistant (PDA), which after pushing the knob 'Calculate risk!' provides an estimate of the risk of persistent symptoms and sick leave for a patient with shoulder pain. In a latter stage the rule will have to change into a decision rule, providing recommendations regarding treatment that may be most beneficial for that particular patient.

Prediction of outcome of shoulder pain

Anamnesis

1	How many weeks ago did your shoulder complaints start?	● <6 weeks	● 6-12 weeks	● >3 months
2	Did you have to stay home from work because of shoulder pain in the past 2 months?	● 0 weeks	● 0-1 weeks	● >1 week
3	Did your complaints start gradually (within a few days)?	● yes	● no	
4	In your opinion, are your complaints caused by overuse or strain doing usual activities?	● yes	● no	
5	Do you often perform repetitive movements in your work or spare time (on at least 2 days/week)?	● yes	● no	
6	Do you also have low back pain?	● yes	● no	
7	Do you have any psychological complaints, such as distress, depression, or anxiety?	● yes	● no	
8	How do you rate your shoulder pain on a scale between 0 and 10? (0=no pain; 10=most severe pain)	● 0	● 1	● 2 ● 3 ● 4 ● 5 ● 6 ● 7 ● 8 ● 9 ● 10

Physical examination

Shoulder pain score	none	little	much	severe
Abduction (active)	●	●	●	●
Abduction (passive)	●	●	●	●
Exorotation (passive)	●	●	●	●
Impingement	●	●	●	●
'Hand in back'	●	●	●	●
'Hand in neck'	●	●	●	●

Neck pain score (passive)	none	little	much	severe
Forward flexion	●	●	●	●
Extension	●	●	●	●
Rotation in neutral position	●	●	●	●
Rotation in flexed position	●	●	●	●
Rotation in extended position	●	●	●	●
Lateral bending	●	●	●	●

Calculate risk!

Risk of persistent symptoms at 6 weeks	%
Risk of persistent symptoms at 6 months	%
Risk of sick leave during 6 months	%

Reference List

1. Scholten-Peeters GG, Verhagen AP, Bekkering GE, van der Windt DA., Barnsley L, Oostendorp RA, Hendriks EJ. Prognostic factors of whiplash-associated disorders: a systematic review of prospective cohort studies. Pain 2003;104:303-22.
2. The Dutch College of General Practitioners (NHG) Practice Guideline. Accessed at http://nhg.artsennet.nl/upload/104/guidelines2/E08.htm on 1 september 2003.
3. Winters JC, Jongh AC, van der Windt DA, Jonquiere M, Winter AF, van der Heijden GJ et al. Practice Guideline "Shoulder complaints' [In Dutch: NHG-Standaard Schouderklachten]. Huisarts en wetenschap 1999;42:222-31.
4. Cats-Baril WL, Frymeyer JW. The economics of spinal disorders. In: Frymeyer JW, editor. The adult spine: principles and practice. New York: Raven Press, 1991.
5. Van der Windt DAWM, Koes BWK, Boeke AJP, et al. Shoulder disorders in general practice: prognostic indicators of outcome. Br J Gen Pract 1996;46:519-23.
6. Croft P, Pope D, Silman A. The clinical course of shoulder pain: prospective cohort study in primary care. BMJ 1996;313:601-2.
7. Winters JC, Sobel JS, Groenier KH, et al. The long-term course of shoulder complaints: a prospective study in general practice. Rheumatology 1999;38:160-3.
8. Okkes IM, Oskam SK, Lamberts H. Van klacht naar diagnose: Episodegegevens uit de huisartspraktijk. Bussum. Coutinho, 1998.
9. Van der Windt DAWM, Koes BWK, Jong BAD, Bouter LM. Shoulder disorders in General Practice: incidence, patients characteristics, and management. Annals of Rheumatic Diseases 1995;54:959-64.
10. Bot SD, Van der Waal JM, Terwee CB, Van der Windt DAWM, Schellevis FG, Bouter LM et al. Incidence and prevalence of complaints of the neck and upper extremity in general practice. Ann Rheum Dis 2005;64:118-23.
11. Miettinen OS. The need for randomization in the study of intended effects. Stat Med. 1983;2:267-71.
12. Bergman GJ, Winters JC, Groenier KH, Pool JJ, Meyboom-de Jong B, Postema K et al. Manipulative therapy in addition to usual medical care for patients with shoulder dysfunction and pain: a randomized, controlled trial. Ann Intern Med 2004;141:432-9.
13. Van der Windt DAWM, Koes BWK, Deville W, Boeke AJ, De Jong BA, Bouter LM. Effectiveness of corticosteroid injections versus physiotherapy for treatment of painful stiff shoulder in primary care: randomised trial. BMJ. 1998;317:1292-96.
14. Kuijpers T, Van der Windt DAWM, Van der Heijden GJGM, Bouter LM. Systematic review of prognostic cohort studies on shoulder disorders. Pain 2004;109:420-431.
15. Harrell Jr. 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
16. Efron B, Tibshirani R. An introduction to the bootstrap. Monographs on statistics and applied probability. New York: NY: Chapman and Hall; 1993.
17. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med 1999;130:515-24.

18. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000;19:453-73.
19. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Validity of prognostic models: when is a model clinically useful? Semin Urol Oncol 2002;20:96-107.
20. Gilks WR RSSD. Markov Chain Monte Carlo in practice. New York: Chapman & Hall, 1996.
21. Vergouwe Y, Steyerberg EW, Eijkemans M, Habbema J. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. J Clin Epidemiol, in press. 2005

Summary

This thesis presents the results of a study which was carried out within the framework of a large research program: The Dutch Shoulder Study (DSS). The DSS is a comprehensive cohort study, carried out between January 2000 and May 2005. The DSS consists of a prognostic cohort study (presented in this thesis) and three randomised controlled trials, which were carried out alongside each other. Shoulder pain is common and has an unfavourable outcome in many patients. In this thesis firstly the literature on the prognosis of shoulder pain was systematically reviewed (Chapter 2). Subsequently, the total costs during 6 months after first consultation generated by patients experiencing shoulder pain were estimated (Chapter 3). Then, the main objectives of this thesis were studied: derive and validate clinical prediction rules for persistent shoulder symptoms at 6 weeks and 6 months after first consultation, and to predict shoulder pain related sick leave during 6 months following first consultation (Chapter 4-7). Subsequently, the association between psychological factors and the persistence of shoulder pain was studied (Chapter 8). Finally, the results were critically reviewed and final conclusions and recommendations for future research were given (Chapter 9).

In Chapter 2 the literature was systematically scored and 16 studies focusing on the prognosis of shoulder disorders were identified. The methodological quality of these 16 studies was assessed. Six of these were considered to be of relatively 'high quality'. There was a wide variety among the studies in length of follow-up, study population, evaluated prognostic factors, type of outcome measure and method of analysis. Due to this large heterogeneity, we refrained from statistical pooling. Instead, we used a best-evidence synthesis. There was strong evidence that high pain intensity predicted a poorer outcome in primary care populations and that middle age (45-54) was associated with poor outcome in occupational populations. There was moderate evidence that a long duration of complaints, and high disability score at baseline predicted a poorer outcome in primary care. These results need to be interpreted with caution because of the small number of studies on which these conclusions are based, and the large heterogeneity among studies regarding follow-up, outcome measures, and analysis.

In Chapter 3 shoulder pain related costs during the 6 months after first consultation in general practice were determined. Information on the costs associated with health care use and loss of productivity in patients with shoulder pain is very scarce. A prospective cohort study with 6 months of follow-up was conducted among 587 patients with a new episode of shoulder pain. Data on costs were collected by means of a cost diary during 6 months. The mean consumption of direct health and non-health related care was low. The total costs in the 6 months after first consultation in primary care, mostly generated by a small part of the population, were not very high (€ 689). Almost 50% of this concerned indirect costs, caused by sick leave from paid work. A small proportion (12%) of the population generated 74% of the total costs. However, after 6 months 46% of the patients still reported persistent symptoms. More extensive research with a longer follow-up is needed. It is important to include patients with fractures, dislocation, or previous surgery to accurately estimate the total costs of illness for patients consulting with shoulder pain in primary care. These patients were not included in our inception cohort, but possibly generate substantial costs.

In Chapters 4 and 5 clinical prediction rules for persistent shoulder symptoms at 6 weeks and 6 months after first consultation were derived and validated. To develop a clinical prediction rule for calculating the absolute risk of persistent shoulder symptoms for individual patients, a prospective cohort study with 6 months follow-up was conducted among 587 patients with a new episode of shoulder pain. Potential predictors included the results of a physical examination, sociodemographic variables, disease characteristics (duration of symptoms, pain intensity, disability, comorbidity), physical activity, physical workload, and psychosocial factors. The main outcome measure was persistent symptoms at 6 weeks and 6 months as perceived by the patient. Response rates to the follow-up questionnaires were 83% at 6 weeks and 92% at 6 months. The following factors were associated with persistent symptoms at 6 weeks in a multivariable model: long duration of symptoms, gradual onset of pain, high pain intensity, concomitant psychological complaints, repetitive

movements, and increased neck pain score at physical examination. Persistent symptoms after 6 months were predicted by a long duration of symptoms at first consultation, gradual onset of pain, high pain intensity, concomitant low back pain, and increased shoulder pain score at physical examination. The discriminative validity was satisfactory with area under the curves of 0.74 (95% CI 0.70; 0.79) at 6 weeks and 0.67 (95% CI 0.63; 0.71) at 6 months.

The generalisability of these prediction rules was evaluated by applying them to a different but comparable population of patients with a new episode of shoulder pain consulting their general practitioner. The control groups of the trials of The Dutch Shoulder Study were merged (n=212), and used to validate the prediction rules. Generalisability of the prediction rules was tested by studying calibration and discrimination in the validation cohort. The prediction rule for short term outcome (6 weeks) showed reasonable calibration and discriminative ability in the validation cohort. The area under the ROC curve (AUC) was 0.72 compared to 0.74 in the derivation cohort. The prediction rule for long term outcome (6 months) performed less well. Discriminative ability (AUC) decreased to 0.56 in the validation cohort compared to 0.67 in the derivation cohort. The prediction rule for the short term prognosis of shoulder pain in general practice showed good generalisability, whereas the prediction rule for long term outcome showed poor generalisability. Hence, it seems difficult to make accurate predictions of the long term outcome of shoulder pain in general practice.

In Chapters 6 and 7 a prediction rule for shoulder pain related sick leave during 6 months following first consultation was derived and validated. To develop the clinical prediction rule for calculating the risk of shoulder pain related sick leave for individual workers, 350 workers with a new episode of shoulder pain from the prognostic cohort study were used. Potential predictors included the results of a physical examination, sociodemographic variables, disease characteristics (duration of symptoms, sick leave in the 2 months prior to consultation, pain intensity, disability, comorbidity), physical activity, physical workload, and (work related) psychosocial factors. The main outcome measure was sick leave during 6

months following first consultation in general practice. Response rate to the follow-up questionnaire at 6 months was 85%. During the 6 months after first consultation 30% of the workers reported sick leave. 16% reported 10 days sick leave or more. Sick leave during this period was predicted in a multivariable model by a longer duration of sick leave prior to consultation, more shoulder pain, strain (overuse) as a result of usual activities, and concomitant psychological complaints. The discriminative ability of the prediction model was satisfactory with an area under the curve of 0.70 (95% CI 0.64; 0.76). Although 30% of all workers with shoulder disorders reported sick leave during follow-up, the duration of sick leave was limited to a few days in most workers.

The generalisability of this prediction rule was evaluated by testing it in two other populations of workers consulting for a new episode of shoulder pain in primary care. Again, the control groups of the trials were merged. All workers were selected (n=128), and used to validate the prediction rule. Besides this population, the recently conducted Musculoskeletal Disorder Study (BAS) (n=224) was used to validate the prediction rule. Generalisability of the prediction rule was tested by studying calibration and discrimination in the validation cohorts. The prediction rule showed reasonable calibration in both validation cohorts. The discriminative ability, with an area under the ROC curve (AUC) of 0.70 in the derivation cohort was stable in the BAS cohort (AUC 0.71). In the control groups of the three RCTs of the DSS the discriminative ability decreased to an AUC of 0.66. The prediction rule for shoulder pain related sick leave during 6 months following first consultation in primary care showed adequate generalisability to another population of workers with shoulder pain participating in an observational cohort study. In the control groups of the three RCTs the prediction rule performed less well.

Chapter 8 studied the influence of psychological factors on the risk of persistent symptoms and disability at three months after consultation in patients with shoulder pain or low back pain in general practice. Psychological factors are assumed to be of prognostic importance and to predict the transition from acute to persistent pain. The influence of

psychological factors has mainly been studied in low back pain, but may be equally important in other types of pain. Patients presenting in general practice with a new episode of shoulder pain in our prognostic cohort study were compared with patients of prospective cohort study on (sub)acute low back pain. In both patient groups psychological factors (catastrophizing thoughts, distress, somatization, and fear-avoidance beliefs) were measured at baseline. Primary outcome measures after three months were 1) persistent symptoms, and 2) less than 30% reduction in functional disability. Logistic regression analyses were used to study the association between baseline scores on psychological factors and outcome. Interaction with symptom duration at baseline was studied for each of the psychological factors. A total of 587 patients with shoulder pain and 171 patients with low back pain were enrolled in the study. Drop-out rate at three months was 12% in patients with shoulder pain and 4% in patients with low back pain. In patients with shoulder pain most associations of psychological factors with outcome were weak and not statistically significant. In patients with (sub)acute low back pain catastrophizing thoughts and somatization were strongly and significantly associated with a higher risk of persistent symptoms and disability at follow-up. Psychological factors, in particular somatization and catastrophizing thoughts, are more strongly associated with persistent pain and disability in patients who consult their general practitioner for low back pain than in patients with shoulder pain. This seems to indicate that different mechanisms may explain the progression to persistent symptoms among patients with different types of pain in primary care. Additional research is needed to confirm these findings and further explore the role of psychological factors in the development of chronic pain problems.

Finally, in Chapter 9 the results of this thesis are critically reviewed and conclusions are presented. An important message of this study is that the strongest predictors of outcome were disease characteristics, results of physical examination, a single work-load factor, and a single yes or no question regarding concomitant psychological complaints. We think that the presented prediction rules for short term outcome and sick leave are

promising enough to test their clinical usefulness in a daily practice setting. The performance of the prediction rule for long term outcome (persistent symptoms at 6 months) in a daily practice situation is questionable, given its poor performance in the derivation and validation cohort.

Samenvatting

Schouderklachten komen veel voor. Herstel van de klachten duurt lang bij veel mensen. Bij ongeveer 50% van de patiënten zijn de schouderklachten na een half jaar nog steeds aanwezig. In dit proefschrift worden scorekaarten gepresenteerd waarmee de huisarts een voorspelling kan doen over de prognose van een patiënt met schouderklachten.

Hoofdstuk 2 bevat een systematische beoordeling op kwaliteit van alle beschikbare wetenschappelijke literatuur over de prognose van schouderklachten. De kosten die met schouderklachten gepaard gaan zijn geschat in Hoofdstuk 3. In Hoofdstuk 4 tot en met Hoofdstuk 7 worden de scorekaarten gepresenteerd waarmee de huisarts de kans op het voortduren van de schouderklachten na 6 weken en na 6 maanden kan berekenen, samen met de kans op ziekteverzuim in de 6 maanden na het eerste consult. In Hoofdstuk 8 wordt de relatie tussen psychosociale variabelen en het voortduren van schouderklachten beschreven. Uiteindelijk worden in Hoofdstuk 9 de resultaten van dit proefschrift kritisch onder de loep genomen, conclusies getrokken en aanbevelingen voor toekomstig onderzoek gegeven.

In Hoofdstuk 2 zijn na het systematisch doorzoeken van de literatuur 16 onderzoeken gevonden die betrekking hebben op de prognose van schouderklachten. Zes van deze onderzoeken waren van relatief hoge methodologische kwaliteit. De studies bleken nogal te verschillen met betrekking tot de lengte van de follow-up, kenmerken van de onderzoeksgroep, onderzochte prognostische factoren, type uitkomstmaten en statistische analyse. Om conclusies te kunnen trekken is er een analyse uitgevoerd waarbij de kwaliteit en de resultaten van elk onderzoek gewogen zijn. Er bleek sterk bewijs te bestaan dat patiënten met veel pijn bij het eerste bezoek aan de huisarts een minder gunstige prognose hadden dan patiënten met milde pijnklachten. Bij werknemers met schouderklachten bleek een leeftijd tussen de 45 en 54 jaar een ongunstige invloed te hebben op de prognose. Er was verder matig bewijs dat een lange klachtenduur voorafgaande aan het eerste bezoek aan de huisarts en het hebben van veel beperkingen in dagelijkse activiteiten een ongunstige invloed hebben op

de prognose van schouderklachten. Deze conclusies moeten echter met de nodige terughoudendheid worden bekeken, omdat er nog maar weinig onderzoek is uitgevoerd en er grote variatie is in de kwaliteit en de opzet van de beschikbare onderzoeken.

Hoofdstuk 3 omvat de kosten die gepaard gaan met schouderklachten in de 6 maanden na het eerste bezoek aan de huisarts. Informatie over kosten van de medische zorg en verlies van productiviteit ten gevolge van schouderklachten (werkverzuim en arbeidsongeschiktheid) is op dit moment schaars. Om de kosten te schatten hebben 587 patiënten gedurende 6 maanden een kostendagboek bijgehouden. De totale kosten in 6 maanden na het eerste bezoek aan de huisarts waren niet erg hoog (gemiddeld € 689 per patiënt). Bijna 50% van dit bedrag is veroorzaakt door indirecte kosten als gevolg van ziekteverzuim. Slechts een klein deel van de 587 patiënten (12%) was verantwoordelijk voor het grootste deel (74%) van de totale kosten. Dit is opvallend omdat 46% van alle patiënten na 6 maanden nog steeds schouderklachten rapporteren. In dit onderzoek zijn patiënten met fracturen, dislocaties, of patiënten die chirurgisch zijn behandeld uitgesloten van deelname. Toekomstig onderzoek naar de kosten van schouderklachten zou deze patiëntengroep wel kunnen insluiten en alle deelnemers gedurende langere tijd moeten vervolgen, zodat de totale kosten van schouderklachten in Nederland nauwkeurig kunnen worden geschat.

In Hoofdstuk 4 en 5 worden scorekaarten gepresenteerd die zijn ontwikkeld en getest voor het berekenen van de kans op het voortduren van schouderklachten na 6 weken en na 6 maanden. Aan dit onderzoek werkten 587 patiënten mee die de huisarts voor schouderklachten consulteerden. Kort na het eerste consult werden een groot aantal factoren gemeten die mogelijk van invloed zijn op de prognose van schouderklachten. De patiënten ontvingen 6 weken en 6 maanden later een vragenlijst over het beloop van hun klachten. De factoren die het voortduren van schouderklachten na 6 weken bleken te voorspellen waren: lange klachtenduur voorafgaand aan het eerste consult bij de huisarts, het

geleidelijk ontstaan van de klachten, hoge schouderpijn score, bijkomende psychische klachten, het uitvoeren van herhaalde bewegingen en veel nekpijn tijdens het lichamelijk onderzoek. Factoren die het voortduren van de klachten na 6 maanden voorspelden waren: een lange klachtenduur, het geleidelijk ontstaan van de klachten, hoge schouderpijn score, bijkomende lage rugklachten en veel schouderpijn tijdens het lichamelijk onderzoek. De scorekaarten bleken redelijk goed in staat mensen met een hoge kans te onderscheiden van patiënten met een lage kans op het voortduren van hun schouderklachten.

Na het testen van de scorekaarten in een andere groep van patiënten met schouderklachten bleek de scorekaart voor 6 weken nog steeds redelijk goed te voorspellen. De scorekaart voor het voorspellen van klachten na 6 maanden presteerde stukken minder en voegde niet veel informatie toe aan een voorspelling op basis van toeval. Het blijkt dus moeilijk uitspraken te doen over de lange termijnprognose van een patiënt met schouderklachten met behulp van een scorekaart die vlak na het eerste consult bij de huisarts wordt ingevuld.

Hoofdstuk 6 en 7 behelst een andere scorekaart die is ontwikkeld en getest om aan schouderklachten gerelateerd ziekteverzuim te voorspellen in 6 maanden na het eerste bezoek aan de huisarts. Hiervoor zijn 350 werknemers geselecteerd en gedurende 6 maanden na het eerste consult gevolgd. Gedurende de 6 maanden na het eerste consult rapporteerde 30% van de werkers werkverzuim, van wie 16% 10 dagen of meer verzuimden. Factoren die ziekteverzuim voorspellen waren: langere periode van ziekteverzuim voorafgaand aan het eerste bezoek aan de huisarts, hoge pijnscore, oorzaak van de schouderklachten is overbelasting door gebruikelijke activiteiten en bijkomende psychische klachten. De scorekaart was redelijk goed in staat mensen met een hoge kans te onderscheiden van patiënten met een lage kans op ziekteverzuim.

De scorekaart is vervolgens getest in twee andere groepen werknemers met schouderklachten. De scorekaart presteerde net zo goed in een onderzoeksgroep met vrijwel dezelfde kenmerken als de groep waarin de kaart ontwikkeld was. De prestaties van de scorekaart waren iets minder

goed in de tweede onderzoeksgroep, die bestond uit de controlegroepen van enkele onderzoeken naar de effectiviteit van de behandeling van schouderklachten.

In Hoofdstuk 8 is de invloed bestudeerd van psychologische factoren op het voortduren van klachten en beperkingen in dagelijkse activiteiten bij patiënten met schouderklachten in vergelijking met patiënten met (sub)acute lage rugklachten. Psychologische factoren worden verondersteld van invloed te zijn op de ontwikkeling van acute naar chronische pijn. Bij schouderpatiënten zijn alleen zwakke associaties gevonden tussen psychologische factoren (gemeten direct na het eerste bezoek aan de huisarts) en klachten of beperkingen na 3 maanden, die niet statistisch significant waren. Bij patiënten met lage rugklachten waren 'catastroferen' en 'somatisatie' sterk en significant geassocieerd met een hogere kans op klachten en beperkingen na 3 maanden. Het lijkt er dus op dat verschillende mechanismen een rol spelen in de ontwikkeling van chronische pijn bij verschillende soorten pijnklachten. Extra onderzoek is nodig om deze bevindingen te bevestigen en om de rol van psychologische factoren in de ontwikkeling van chronische klachten aan het bewegingsapparaat verder te onderzoeken.

Uiteindelijk wordt in de algemene discussie (Hoofdstuk 9) geconcludeerd dat langdurige schouderklachten en ziekteverzuim het sterkst kunnen worden voorspeld met ziektekenmerken, uitkomsten van lichamelijk onderzoek, een enkele werkgerelateerde vraag en een ja/nee vraag over psychische klachten. Het is belangrijk om de klinische toepasbaarheid van de scorekaarten voor de korte termijnprognose (6 weken) en voor ziekteverzuim te testen in de dagelijkse praktijk. De toepasbaarheid van de scorekaart voor de lange termijn (6 maanden) is twijfelachtig, gezien de tegenvallende prestaties wanneer de kaart wordt toegepast in andere groepen patiënten.

Thanks to

Alan Silman • Wim Stalman • Raymond Ostelo • Carl Moons • Monique Frings-Dresen • Bart Koes Faas • Sandra Bot • 103 huisartsen • 587 patienten • Brahim Oughzou • Cor Brasser • Cees de Boer • Wim Kraan • Michel Telkamp • Ellen Visser • Rob de Bie • Marielle Goossens • Jan Winters Arlette Hesselink • Lonneke van de Poll • Petra Jellema • Koen van der Kooij • Marleen Hermens Els Licht • Bettine Schreuders • Ingrid Bakker • Esther van Sluijs • Amika Singh • Hein van Hout Berend Terluin • Els Pronck • Sophal Variosseau • Valentien Blom • Gieter • Antoon & Dorothée Gouw# • De Passie • De Kutzeeuws • Mariska

Curriculum Vitae

Geboren te Oisterwijk	1974
Opleiding Fysiotherapie Hogeschool West-Brabant, Breda	1994-1997
Opleiding Bewegingswetenschappen Universiteit Maastricht	1997-1999
Opleiding Epidemiologie Universiteit Maastricht	1999
Promovendus EMGO-instituut, VU Medisch Centrum	2000-2005
Fysiotherapeut Paramedicum Utrecht	2000-2003
Didactische opleiding IDO, VU Amsterdam	2000
Docent Evidence Based Practice Landelijke verplichte nascholing voor Fysiotherapeuten	2002-2003
Onderzoeker/docent Lectoraat, Amsterdamse Hogeschool voor Paramedische opleidingen	2005
Adviseur Richtlijnen/Indicatoren Kwaliteitsinstituut voor de gezondheidszorg CBO	2005

frikadel