Original article

Scand J Work Environ Health 2007;33(6):440–446

Generalizability of a prediction rule for sick leave due to shoulder pain

by Ton Kuijpers, PhD,[1, 2, 3] Yvonne Vergouwe, PhD,[4] Geert JMG van der Heijden, PhD,[4] Sandra DM Bot, PhD,[1] Jos WR Twisk, PhD,[1, 5] Daniëlle AWM van der Windt, PhD,[1, 6] Lex M Bouter, PhD [1]

Kuijpers T, Vergouwe Y, van der Heijden GJMG, Bot SDM, Twisk JWR, van der Windt DAWM, Bouter LM. Generalizability of a prediction rule for sick leave due to shoulder pain. *Scand J Work Environ Health*. 2007;33(6):440–446.

Objectives Recently, a rule was developed to predict sick leave related to shoulder pain during a period of 6 months after patients have consulted a general practitioner for a new episode of shoulder pain. The objective was to evaluate the generalizability of this prediction rule by testing it in two other populations of workers who had gone for a consultation in primary care for a new episode of shoulder pain.

Methods The prediction rule was derived in a prognostic cohort study (N=350). The outcome was sick leave related to shoulder pain during 6 months following the first consultation. The rule was tested on merged control groups from three trials on shoulder pain (N=128). In addition to this population, a recently conducted study on musculoskeletal disorders (N=224) was used to validate the prediction rule. The generalizability 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 receiver operating characteristic (ROC) curve (AUC) of 0.70 in the derivation cohort was stable in the cohort of the musculoskeletal disorder study (AUC 0.71). In the control groups of the three randomized controlled trials of a Dutch shoulder study, the discriminative ability decreased to an AUC of 0.66.

Conclusions The prediction rule for sick leave related to shoulder pain in a 6-month period following the first consultation in primary care showed adequate generalizability to another population of workers with shoulder pain participating in an observational cohort study. In the control groups of the three randomized controlled trials the prediction rule performed less well.

Key terms absenteeism; cohort study; musculoskeletal disorder; prognosis.

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

The early identification (risk stratification) of patients with a high risk of sick leave due to shoulder pain may enable timely intervention and prevent sick leave and the concomitant high costs these patients generate. We developed a clinical prediction rule consisting of the following four easily measurable prognostic factors: sick leave at baseline in the preceding 2 months (0/0–1/≥1 weeks), shoulder pain (0–10), strain or overuse due to usual activities as a precipitating cause (yes, no), and concomitant psychological complaints (yes, no) (table 1). Points are given for each of the four prognostic factors, and a total score is calculated.

The total score corresponds with the estimated risk for sick leave due to shoulder pain during the 6 months following a first consultation for patients with a new episode of shoulder pain in primary care. The performance

1 Institute for Research in Extramural Medicine, VU University Medical Center, Amsterdam, Netherlands.
2 Department of Allied Health Care Research, Amsterdam School of Allied Health Care Education, Amsterdam, Netherlands.
3 Care and Public Health Research Institute, Maastricht University, Maastricht, Netherlands.
4 Julius Center for Health Sciences and Primary Care, University Medical Center, Utrecht, The Netherlands.
5 Department of Clinical Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, Netherlands.
6 Primary Care Musculoskeletal Research Centre, Keele University, Keele Staffordshire, United Kingdom.

Correspondence to: Dr T Kuijpers, Dutch Institute for Healthcare Improvement (CBO), Churchilllaan 11, NL-3502 LB Utrecht, Netherlands. [E-mail: t.kuijpers@cbo.nl]

Table 1. Prognostic score chart for predicting sick leave during 6 months following a first consultation. 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 problems)].[a]

			Total score	Risk (%)
Sick leave in the preceding 2 months				
None	0	…	≤1	10–20
0–1 week	2	…	2–3	20–30
>1 week	3	…	4–5	30–40
Intensity of shoulder pain (0–10)			6–7	40–50
0–3 points	0	…	8	50–60
4–6 points	2	…	9–10	60–70
7–10 points	3	…	11–12	70–80
Perceived cause: strain or overuse during regular activities	3	…	13–15	80–90
Reported psychological problems (anxiety, distress, depression)	6	… +		
Total score		…		

[a] 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". With the use of the table next to the score chart, the risk (%) of sick leave for an individual patient can be determined on the basis of the total score.

(ie, calibration and discrimination) of the prediction rule was evaluated in the development study (12). Calibration refers to the extent to which 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 for sick leave and a patient who will not have to stay off work because of shoulder pain.

Before the implementation of the prediction rule in clinical practice was considered, its generalizability needed to be tested (13–15). Generalizability refers to the performance of patients drawn from a different but comparable population (13). Our objective was to evaluate the performance of our clinical prediction rule for sick leave due to shoulder pain in two different cohorts of patients with shoulder pain in a primary care setting.

Study population and methods

In this study, we evaluated the generalizability of the derived prediction rule from a Dutch shoulder study in a subgroup of other patients from this cohort and among participants of another prospective cohort study in general practice, the musculoskeletal disorder study (MDS).

Dutch shoulder study

The Dutch shoulder study (DSS) was a comprehensive cohort study carried out between January 2000 and May 2005. It consisted of one prognostic cohort study and three randomized controlled trials, which were carried out in parallel (see figure 1). Between January 2001 and June 2003, 103 general practitioners recruited patients at their first consultation for a new episode of shoulder complaints in three geographic areas in the Netherlands (Amsterdam, Groningen, and Maastricht).

All of the patients in the Dutch shoulder study had to meet the following general inclusion criteria: was 18 years of age or older; had paid work; had not consulted a general practitioner or received any form of treatment for the afflicted shoulder in the preceding 3 months; and had sufficient knowledge of the Dutch language.

The patients had to meet the specific additional inclusion criteria of the trial in which they were included, if eligible, as follows: dysfunction of the cervicothoracic spine and adjacent ribs with accompanying pain or restricted movement in the *Groningen manipulation study*, >3-month duration of symptoms in the *graded exercise therapy study*, and <3-month duration of symptoms in the *education and activation program*.

In both the Dutch shoulder study and the musculoskeletal disorder study, those with severe physical or psychological conditions (ie, fractures or dislocation in the shoulder region, rheumatic disease, neoplasm, neurological or vascular disorders, dementia) were excluded. 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 generalizability of the rule. For the current study only the patients who reported paid work were used.

The Groningen manipulation study (16–17) evaluated the effectiveness of manipulative therapy for the shoulder girdle in addition to usual care. In the two other trials, graded exercise therapy (18) and an education and activation program (19) were used. The participants

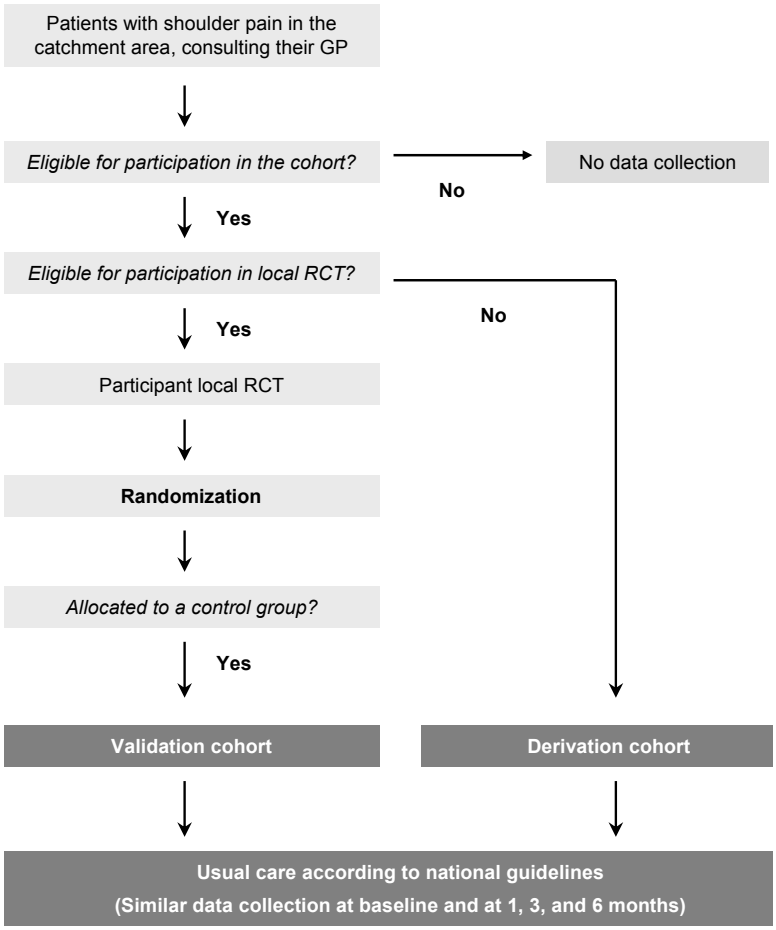


Figure 1. Design of the Dutch shoulder study. (GP = general practitioner, RCT = randomized controlled trial)

were evenly allocated to the treatment groups on the basis of a random list prepared by an independent statistician not involved in recruiting the patients. A research assistant opened preprepared, numbered, opaque envelopes, which were sealed and contained the treatment allocation codes (16–19). The patients in the control groups of the trials received usual care, similar to that of the patients in the cohort study.

Baseline and follow-up assessments for all of the patients were identical in the Dutch shoulder study. The outcome was measured with the use of postal questionnaires at 6 weeks, 3 months, and 6 months.

Musculoskeletal disorder study

The musculoskeletal disorder study (MDS) is a large observational cohort study that was conducted in 61 general practices (97 general practitioners) (20–21). The physicians recruited patients who consulted them due to a new episode of musculoskeletal pain. For our generalizability study, we selected patients who came for a consultation for shoulder pain and who had paid work at baseline. The selection criteria were comparable with those of the Dutch shoulder study. Follow-up questionnaires were sent after 3, 6, and 12 months.

Prediction rule

The rule predicted sick leave due to shoulder pain (yes = ≥1 day, no = 0 days) during 6 months after the first consultation with a physician about shoulder pain and was developed using information from the 350 patients of the derivation cohort who reported paid work at baseline. Sociodemographic variables, disease characteristics (ie, pain intensity, disability, duration of complaints, sick leave in the 2 months prior to the 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 (eg, 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 0.70 [95% confidence interval (95% CI) 0.64–0.76]. In table 1 on page 441 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 process, the patients were grouped into quintiles according to their predicted probability of sick leave. The prevalence of the end point within each quintile equaled 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 indicated no discrimination above chance, whereas an AUC of 1.0 indicated 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 AUC attainable for the validation cohorts. With the use of the predicted risks of the patients in the validation cohorts, outcomes were generated with Monte Carlo simulation (22–23). The simulation mimics a situation in which the model is perfectly calibrated, thereby showing the extent to which poor calibration can affect the discriminative performance of the model. 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 showed which of the different elements of the rule were 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. The 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 musculoskeletal disorder study were less often male (42% versus 55%), more often reported strain or overuse due to usual activities as a precipitating cause

Table 2. Description of the baseline characteristics of the workers with shoulder pain in the derivation and validation cohorts. (DSS= Dutch shoulder study, MDS = musculoskeletal disorder study, SD= standard deviation, IQR = interquartile range)

	Derivation cohort				Validation cohort							
	DSS (N=350)				DSS controls (N=128)				MDS (N=224)			
	N	%	Mean	SD	N	%	Mean	SD	N	%	Mean	SD
Demographic characteristics												
Age (years)	45	11	·	·	46	9	·	·	44	10	·	·
Gender (male)	193	55	·	·	63	50	·	·	95	42	·	·
Disease characteristics												
Duration of complaints >3 months	134	38	·	·	65	51	·	·	90	40	·	·
Sick leave at baseline in preceding 2 months [a]												
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 or overuse, usual activities [a]	99	28	·	·	40	31	·	·	104	46	·	·
Shoulder complaints in the past	199	57	·	·	79	62	·	·	114	50	·	·
Concomitant psychological complaints [a]	27	8	·	·	10	8	·	·	37	17	·	·
Concomitant musculoskeletal complaints												
Neck or high back-pain	119	34	·	·	51	40	·	·	128	58	·	·
Low-back pain	61	17	·	·	34	27	·	·	77	34	·	·
Shoulder pain (0–10) [a]	·	·	4.5	2.3	·	·	5.2	2.3	·	·	4.9	2.2
Work-related psychosocial factors [b]												
Quantitative job demands (4–20)	·	·	12.8	2.7	·	·	12.6	2.6	·	·	13.1	3.2
Decision authority (3–12)	·	·	9.4	1.8	·	·	9.6	1.9	·	·	9.1	2.1
Co-worker support (4–16)	·	·	12.3	2.0	·	·	12.3	1.9	·	·	12.3	1.7

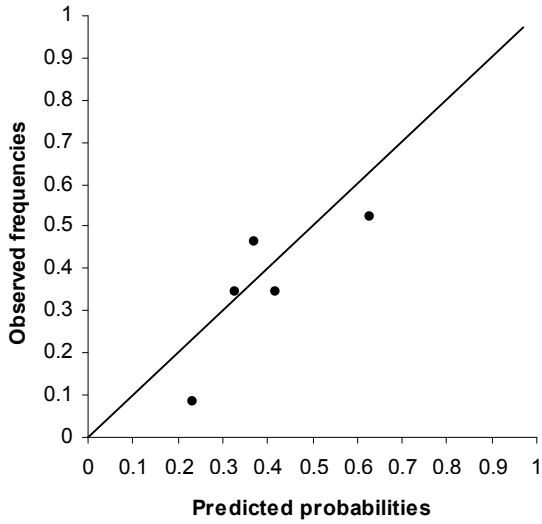
[a] Variables that are in the rule predicting sick leave due to shoulder pain.

[b] A score between 10 and 15 points reflects fair quantitative job demands; a score of >9 points reflects high decision authority; a score of >12 points reflects high co-worker support.

(46% versus 28%), and more often reported concomitant musculoskeletal complaints of the neck or high back (58% versus 34%) and low back (34% versus 17%). The patients in the derivation and validation cohorts reported similar percentages of sick leave for the first 6 months after the first consultation (30% for the DSS derivation cohort, 34% for the DSS control groups, and 32% for the MDS cohort).

In the MDS cohort, 90% of all of the patients completed the 6-month questionnaire. There were no significant differences between the respondents and dropouts with respect to gender, functional disability, or pain, although the dropouts were slightly younger. For the DSS controls, the dropout rate at 6 months was <10%. No information was available in regard to the differences between the respondents and the dropouts.

DSS controls



MDS

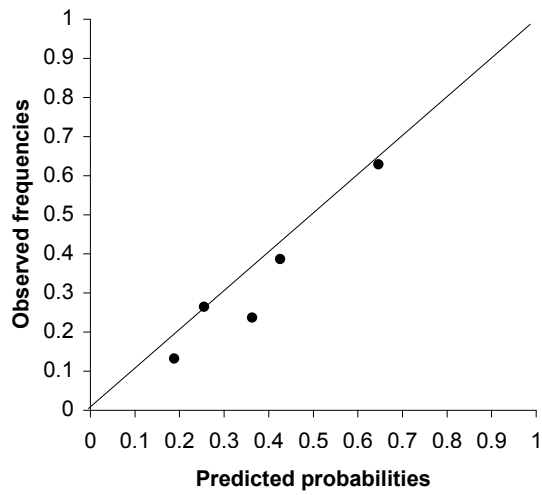


Figure 2. Calibration plots showing the observed frequencies versus the predicted probabilities for sick leave due to shoulder pain during the 6 months following the first consultation in primary care, for the controls of the Dutch shoulder study (DSS) (N=103) and a cohort of the musculoskeletal disorder study (MDS) (N=176). The patients were grouped into quintiles according to their predicted probability of sick leave due to shoulder pain according to the prediction rules. The prevalence of the end point within each quintile represents the observed individual probability.

Performance

Figure 2 shows the calibration of the predictions. For the DSS control groups the predicted risks for sick leave were 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 MDS cohort most of the plotted points were rather close to the 45-degree line, although three prediction categories slightly overestimated the observed probabilities. Again, the mean predicted probability (0.39) was only slightly higher than the overall observed sick leave prevalence (0.34). The discriminative ability (AUC) 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 MDS cohort. The results of the Monte Carlo simulation show that these estimates were close to the maximum attainable AUC (0.66 for the DSS controls and 0.70 for the MDS population), indicating that reductions in discrimination were not explained by poor calibration.

Table 3 shows the multivariate regression coefficients when the rule was applied to the validation cohorts. Shoulder pain was a strong predictor of sick leave in the DSS control groups and in the MDS cohort. In the MDS cohort, sick leave in the 2 months preceding the baseline examination also showed a strong relation to sick leave during the follow-up. The category "0–1 weeks" showed a remarkable negative association with outcome among the DSS controls. A similar opposite association was found for the category "4–6 points, shoulder pain" in the MDS cohort.

Discussion

The performance of the prediction rule for sick leave due to shoulder pain in the DSS control groups showed an unstable calibration and a slightly decreased discriminative ability (AUC of 0.66, compared with an AUC of 0.70 in the derivation cohort). The prediction rule calibrated better for the MDS population and showed a stable discriminative ability (AUC 0.71).

The validation of a prediction rule is preferably started with populations that are very similar to the derivation cohort. In the next step, the rule is tested in populations that show more differences from the

derivation cohort in terms of, for example, baseline characteristics or setting (13). The DSS was designed specifically to design and validate a clinical prediction rule for shoulder pain. The patients in both the derivation and validation cohorts (DSS control groups) were recruited by the same group of general practitioners, and they received similar treatment. The measurements were exactly the same and were carried out by the same research teams. This was an optimal setting in which to start testing the generalizability of the prediction rule. Although there was considerable similarity in the sampling frame and data collection between the DSS controls and the MDS participants, a different group of general practitioners participated in the MDS, the treatment may have been somewhat different, and the data collection showed some differences. The MDS cohort was therefore a good cohort for use in the second step of the validation process. At the time of our study, no other good cohorts were available.

Patients that originate from clinical trials usually form a more selective population than patients participating in observational studies. Although the trials in the DSS used fairly broad selection criteria, this difference may have influenced the performances of the prediction rule with respect to the DSS controls. For example, the DSS controls included more patients with a long symptom duration, which is an important prognostic factor. In addition, it may be possible that the patients who participated in the trials had different characteristics that were not measured in our study, for example, regarding treatment expectations or preferences.

Finally, the unstable calibration with the DSS control groups may also have partly been a result of the small numbers (N=103) in this validation cohort. Other studies have also shown that the calibration of prediction rules can be unstable when applied to 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 among 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 have been a result of differences in the baseline characteristics between the derivation and this validation cohort.

In the MDS cohort, shoulder pain and sick leave at baseline in the preceding 2 months showed substantial and significant regression coefficients (table 3). This phenomenon, combined with higher regression coefficients for strain or overuse due to usual activities and concomitant psychological complaints in the MDS cohort in comparison with the DSS controls, may have resulted in a better performance of the prediction rule for the MDS cohort. The substantial baseline differences between the derivation cohort and the MDS, in regard to factors that 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 of 0.70 in the MDS cohort strengthens our findings of adequate discriminative ability in this cohort.

Table 3. Multivariate regression coefficients for sick leave due to shoulder pain during 6 months after the first consultation in the validation in comparison with the coefficients obtained from the derivation cohort. The β values were derived from a multiple logistic regression analysis and were shrunk using bootstrapping techniques. The β values for the validation cohorts were computed by entering the four predictors of the rule simultaneously into a multiple logistic regression analysis, demonstrating the strength of the association between the predictors and sick leave in the validation cohorts in comparison with the derivation cohort. (DSS = Dutch shoulder study, MDS = musculoskeletal disorder study, β= regression coefficient, 95% CI = 95% confidence interval)

Predictor	Derivation		Validation			
	DSS (N=286)		DSS controls (N=103)		MDS (N=176)	
	β	95% CI	β	95% CI	β	95% CI
Sick leave at baseline (in preceding 2 months)						
0–weeks [a]						
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 [a]						
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 or 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

[a] Reference category.

We developed a rule to predict sick leave due to shoulder pain during the first 6 months after a 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. Therefore, future research should also evaluate the methodological transportability of the prediction rule [ie, performance when data are collected with alternative methods (13) in a new sample of workers], and, perhaps most importantly, the clinical usefulness of the instrument should be established. In other words, 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 sick leave due to shoulder pain during the first 6 months after a first consultation showed disappointing generalizability for the DSS controls but adequate generalizability for the MDS observational cohort.

To enhance the performance of prediction rules in the future, rules should be updated and amended when new and larger cohorts become available. Validation should preferably be carried out prospectively, and—when sick leave is being predicted—in an occupational setting. Alternatively, separate models can be designed for patients with either acute or chronic shoulder problems, as different factors may be important in predicting outcome in these different patient groups. This step may further enhance the predictive performance of prediction rules for shoulder pain.

Acknowledgments

This study was supported by a grant (904–65–09) from the Netherlands Organisation for Scientific Research (NWO), The Hague, Netherlands.

Reference

1. 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.
2. Van der Heijden GJ. 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 DA, 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, Schellevis FG. Klachten en aandoeningen in de bevolking en in de huisartsenpraktijk [Complaints and disorders in the population and general practice]. Utrecht/Bilthoven (Netherlands): Netherlands Institute for Health Services Research; 2005.
8. 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.
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 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.
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 DA, Van der Heijden GJ, Twisk J, Vergouwe Y, Bouter LM. A prediction rule for shoulder pain related sick leave: a prospective cohort study. BMC Musculoskelet Disord. 2006;7:97.
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 GJ, 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 DA, Scholten RJ, Bouter LM, et al. Predictors of outcome in neck and shoulder complaints: a cohort study in general practice. Spine. 2005;30:459–70.
22. Gilks WR RSSD. Markov Chain Monte Carlo in practice. New York (NY): Chapman & Hall; 1996. X
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. 2005;58:475–83.
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.

Received for publication: 27 February 2007