



prediction model could aid clinicians, and their patients in making informed decisions on treatment possibilities, taking into account their prognosis and risk for persistence. Currently existing clinical prediction models or risk prediction tools all focus on younger adults<sup>10,18,55,57,58</sup> and often focus on acute back pain alone.<sup>10,18,57</sup> Therefore, developing and validating a clinical prediction model for persistent back pain in an older population is important to identify elderly patients at risk of poor recovery in primary care.

Although prognostic factor research in older adults with back pain has shown that similar demographic, back pain-related, general health, and psychological factors may play a role in nonrecovery,<sup>47,49</sup> the evidence is still scarce. In fact, the existing results are preliminary, because, in 2 existing studies, the focus was on a 3-month follow-up in one<sup>49</sup> and only on univariable associations in the other.<sup>47</sup> The current study aimed to develop a multivariable clinical prediction model for nonrecovery in a well-characterized population-based general practice cohort of older adults seeking care for back pain and to assess its internal validity according to current research guidelines.<sup>53</sup>

## 2. Methods

### 2.1. Design, setting, and patient recruitment

Back Complaints in the Elders (BACE) is a prospective cohort study aiming to explore the course and prognosis of back pain in older people. In total, 103 general practitioners (GPs) from the southwest area of the Netherlands participated in this study, that is part of the international BACE consortium.<sup>52</sup> Six-hundred and 75 participants aged > 55 years who consulted their GP with a new episode of back pain in 2009 to 2011 were included. An episode was defined as “new” if the patient had not consulted a GP for the same back pain symptoms in the preceding 6 months; back pain was defined as pain in the region from the top of the shoulder blades to the first sacral vertebra. Patients who were unable to complete the questionnaires (eg, because of an inability to read and write in Dutch or cognitive impairment) and/or undergo the physical examination (eg, patients who use a wheelchair) were excluded. Back Complaints in the Elders was approved by the local medical ethics committee, and a detailed study protocol was previously published.<sup>52</sup> This article is reported according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines for multivariable prediction model research.<sup>9,35</sup> The Prediction model Risk Of Bias Assessment Tool (PROBAST) was used as an additional checklist to minimize the risk of bias of this study.<sup>36,64</sup> According to the Prognosis Research Strategy (PROGRESS) initiative, this article is part of the first phase of prognostic model research: model development, including internal validation.<sup>53</sup>

### 2.2. Data collection

Baseline data were collected using questionnaires, history taking, and physical examination. Baseline measurements included demographics and patients’ characteristics (eg, age, sex, body mass index, education level, smoking, use of pain medication, or care from a physical therapist for current back pain); measurements on the physical domain, such as back pain-related factors (eg, back pain duration and pain intensity), physical functioning, and physical examination items; and

measurements on the psychosocial domain. Health-related quality of life on a physical and mental level was measured with the Short Form-36 (SF-36; physical/mental component summary scales (both range 0-100)).<sup>63</sup>

The BACE cohort included follow-up measurements at 6 weeks, 3, 6, and 9 months and 1, 2, 3, 4, and 5 years. In this study, responses at baseline (n = 670), 6 months (n = 612), and 1 year (n = 591) were used because the response rates up to 1-year follow-up were relatively high, and it is a relevant time period for this condition. For the development of the prediction model, the baseline data were used as potential predictors; the combined 6-month and 1-year follow-up data were used as outcomes.

### 2.2.1. Outcome measures

There are many potential definitions for nonrecovery in back pain, but a general and widely used definition of nonrecovery has not been established yet.<sup>29</sup> Nevertheless, various studies in patients with back pain exhibited that pain and physical functioning are the 2 most important outcomes, reasons to seek care, and part of a broad definition of recovery.<sup>1,7,28,33,65</sup> Therefore, the following outcome measures were used: the 11-point numeric rating scale (NRS, ranging from 0 “no pain” to 10 “worst pain ever”)<sup>61</sup> to measure pain intensity and the 24-item Roland Morris Disability Questionnaire (RMDQ, ranging from 0 “no disabilities” to 24 “maximal disability”)<sup>45</sup> to measure physical functioning. These 2 instruments also represent core outcome measures in patients with low back pain.<sup>6</sup>

A previous USA-based cohort of older adults with back pain (BOLD cohort) defined the presence of persistent pain and persistent disability using these outcome measures.<sup>47</sup> To facilitate comparability of findings, the same definitions were used in the current study. Nonrecovery as defined by Scheele et al.<sup>49</sup> was added as a third outcome, dichotomizing the 7-point Global Perceived Effect (GPE, ranging from 1 “completely recovered” to 7 “worse than ever”) scale<sup>4</sup> to measure self-perceived recovery. In this case, recovery was defined as a score of 1 or 2 of 7 (either “completely recovered” or “strongly improved”) at both time points. Any other response option (ranging from “somewhat improved” to “worse than ever”) was defined as nonrecovery.

Thus finally, nonrecovery was defined as: (1) persistent back pain: back pain intensity of 3/10 points or higher on the NRS at both 6 and 12 months, (2) persistent disability: disability of 4/24 points or higher on the RMDQ at both 6 and 12 months, and (3) perceived nonrecovery: a score of 3/7 points or higher on the GPE scale at both 6 and 12 months.

### 2.2.2. Sample size calculation and predictor selection

Considering the BACE sample size, the maximum number of predictors in the full model was calculated using the approach of Riley et al.<sup>43</sup> The proportion of nonrecovery outcomes was 48% for pain, 54% for physical functioning, and 52% for GPE. For all 3 outcomes, this resulted in a maximum number of 14 candidate predictors.

Potential predictors were identified as follows: (1) those consistently found to be significantly and consistently associated to outcomes in the previous (low) back pain literature<sup>5,8,11,14-17,19,20,31,39-41,47,49</sup> and (2) those deemed potentially clinically relevant, based on the clinical experience of the BACE project team including GPs and physiother-

Table 1

## Baseline patients' characteristics in the Back Complaints in the Elders prospective cohort.

	All, n = 675	Missing, n (%)
General characteristics and demographics		
Age in y, mean ± SD	66.4 ± 7.6	0 (0)
Male, n (%)	274 (41)	0 (0)
BMI, mean ± SD	27.5 ± 4.7	7 (1)
Low education level, n (%)	279 (41)	7 (1)
Smoking, n (%)	122 (18)	8 (1)
Use of pain medication for current back pain, n (%)	485 (72)	9 (1)
Care from a physical therapist for current back pain, n (%)	299 (44)	17 (3)
Physical domain		
Duration of current back pain episode > 3 mo, n (%)	156 (23)	80 (12)
Average back pain intensity the previous wk (NRS), mean ± SD	5.2 ± 2.7	11 (2)
Physical functioning (RMDQ), mean ± SD	9.9 ± 5.8	16 (2)
At least 1 back pain episode during lifetime, n (%)	579 (86)	9 (1)
Recent episode of back pain (in previous 6 mo), n (%)	487 (72)	10 (2)
Musculoskeletal comorbidity, n (%)	519 (77)	12 (2)
Referred pain to the leg, n (%)	379 (56)	7 (1)
Spinal morning stiffness of the back for > 30 min, n (%)	160 (24)	14 (2)
Pain during spinal rotation, n (%)	259 (38)	7 (1)
Quality of life (SF-36) physical summary scale, mean ± SD	43.2 ± 8.9	7 (1)
Psychosocial domain		
Patients' expectation to recover from back pain, n (%)	291 (43)	17 (3)
Quality of life (SF-36) mental summary scale, mean ± SD	49.6 ± 10.3	7 (1)
Depressive symptomatology (CES-D), mean ± SD	10.0 ± 7.8	57 (8)
Kinesiophobia (FABQ), physical activity subscale, mean ± SD	13.4 ± 5.8	20 (3)
Pain catastrophizing (PCS), mean ± SD	14.1 ± 10.6	33 (5)

BMI, body mass index (kg/m<sup>2</sup>); NRS, numeric rating scale (range 0–10); RMDQ, Roland Morris Disability Questionnaire (range 0–24); SF-36, Short Form-36, physical and mental summary scale (range 0–100); CES-D, Center for Epidemiologic Studies Depression Scale (range 0–60); FABQ, Fear-Avoidance Beliefs Questionnaire, physical activity subscale (range 0–28); PCS, Pain Catastrophizing Scale (range 0–52).

apists, and a discussion meeting with GP trainees.

Forty-five potential predictors that might be of interest were identified in the data set, all part of the questionnaires and the physical examination. X-ray variables were not included, because the clinical prediction model should be easy to use in general practice, and current guidelines do not recommend routine use of imaging unless serious pathology is suspected.<sup>37</sup> After extensive literature review and a meeting with the BACE project team, this number was reduced to 18. Some potential predictors were excluded because they were not a factor known in the literature; others were (clinically) highly related to each other, and one of both measures was chosen. After a discussion meeting with GP trainees, the number was further reduced to the maximum of 14, taking into consideration the potential clinical relevance and usefulness in general practice.

After predictor selection, the 14 candidate predictors were the following: age, sex, chronic duration (current episode duration of >3 months at baseline), back pain intensity in the past week, back pain-related disability, a recent preceding episode of back pain, musculoskeletal comorbidity, referred pain to the leg, spinal morning stiffness, pain during spinal

rotation, patients' expectations to recover, depressive symptomatology, kinesiophobia, and pain catastrophizing.

These candidate predictors were measured as followed (unless previously mentioned): back pain intensity measured with the 11-point NRS,<sup>61</sup> referring to the average pain intensity over the last week, back pain-related disability measured with the 24-item RMDQ,<sup>45</sup> the presence of other musculoskeletal symptoms was explored, patients were asked if they experienced a recent episode of back pain where they did not seek care for (ie, in the previous 6 months before study inclusion, without visiting the GP for these symptoms), radiating pain in the leg(s), spinal morning stiffness for longer than 30 minutes, and pain during spinal rotation of the upper body to the left and/or right (all dichotomous). Depressive symptomatology was measured with the Center for Epidemiologic Studies Depression Scale (CES-D, range 0–60),<sup>42</sup> kinesiophobia with the physical activity subscale of the Fear-Avoidance Beliefs Questionnaire (FABQ, range 0–28),<sup>62</sup> and pain catastrophizing with the Pain Catastrophizing Scale (range 0–52).<sup>54</sup> The patients' expectations to recover was measured using a 5-point Likert scale ranging from 1 "completely pain free" to 5 "more pain than ever."

### 2.3. Statistical analysis

Descriptive statistics were used to describe the baseline patient characteristics. Frequencies were used for categorical data, and mean and standard deviation were provided for continuous data.

Primary analyses were performed in the imputed data sets. Multiple imputation was performed using all predictor and outcome variables to impute missing data. Following guidelines on missing data, 25 imputed data sets were created with 50 iterations because 25% of the cases had missing values (a total of 3% missing values), and predictive mean matching was used.<sup>23</sup> In multi-item questionnaires, the separate item scores were imputed instead of the sum scores, to improve accuracy of the estimates, with the exception of the RMDQ, where missingness was taken into account using an existing scoring algorithm.<sup>30</sup> Convergence plots were generated to check if the imputations were correctly performed.

A univariable analysis of the candidate predictors was performed to assess the relationship between the separate predictor variables and the outcomes. This analysis was performed for informative reasons but not used for predictor selection.<sup>44</sup> The linearity assumption for continuous variables was evaluated using restricted cubic splines, and multicollinearity was evaluated by determining tolerance (>0.2) and the variance inflation factor (<3). When the linearity assumption was not met, splines were added to the models. All candidate predictors were included in the multivariable logistic regression analysis.

A multivariable logistic regression analysis was performed for each of the 3 predefined outcomes in the imputed data sets as primary analysis. The backward selection method was used, with an entry and removal *P* value of 0.157 according to the Akaike's information criterion.<sup>9</sup> This was repeated until the best-fitted models were found. Models' performance was evaluated for each of the 3 models with: (1) the explained variance or overall performance, using Nagelkerke's *R*<sup>2</sup>; (2) calibration, as estimated by the calibration slope, curve, and the Hosmer-Lemeshow test; and (3) discrimination, as measured by the AUC. A *P* value > 0.05 of the Hosmer-Lemeshow test was considered as good calibration. The discriminative performance of the models was considered acceptable if the AUC was ≥0.7 and good if the AUC was ≥0.8.

To correct for overoptimism of the final models, these were internally validated in 250 bootstrapped samples. The estimated slope value was used as a shrinkage factor that was multiplied with the pooled coefficients, and a new intercept value was determined that aligned with the shrunken coefficients. These adjusted linear predictors were reported in the penalized models, and each adjusted models' performance was evaluated with an optimism-adjusted *R*<sup>2</sup> and AUC value. Statistical analyses were performed using SPSS (IBM, version 24), and (R) Studio version 3.4.3, using the *psfmi* 0.5.0 package.<sup>22</sup>

### 2.4. Sensitivity analysis

A complete case analysis was performed for all outcome variables as a sensitivity analyses.

## 3. Results

### 3.1. Baseline characteristics

A total of 675 patients were included. **Table 1** shows the cohorts' baseline characteristics and the amount of missing per variable.

### 3.2. Univariable analysis

**Table 2** shows the univariable analysis of all candidate predictors in the imputed data sets. They were all related to the outcomes. For all 3 outcomes, male sex and a patients' expectation to recover from their back pain are negatively associated with persistence of symptoms.

### 3.3. Model development and performance

All final clinical prediction models are shown in **Table 3**. The 3 models predicting, respectively, persistent back pain, persistent disability, and perceived nonrecovery as an outcome all included age, chronic duration of symptoms, back pain-related disability, a recent episode of back pain, and patients'

**Table 2**  
Univariable associations.

Univariable analysis Candidate predictors	Persistent back pain			Persistent disability			Perceived nonrecovery		
	OR*	95% CI	<i>P</i>	OR*	95% CI	<i>P</i>	OR*	95% CI	<i>P</i>
Age (y)	1.04	1.02–1.06	0.000	1.05	1.03–1.08	0.000	1.04	1.02–1.06	0.001
Gender, male	0.60	0.43–0.84	0.003	0.68	0.49–0.93	0.015	0.64	0.46–0.89	0.008
Chronic duration (>3 mo)	3.40	2.27–5.08	0.000	3.02	1.97–4.62	0.000	3.16	2.09–4.80	0.000
Back pain intensity in the past wk (NRS 0-10)	1.29	1.21–1.38	0.000	1.27	1.19–1.36	0.000	1.20	1.12–1.27	0.000
Disability (RMDQ 0-24)	1.14	1.11–1.18	0.000	1.25	1.21–1.30	0.000	1.11	1.08–1.14	0.000
Recent episode of back pain (previous 6 mo)	4.20	2.81–6.26	0.000	3.70	2.54–5.38	0.000	5.58	3.70–8.42	0.000
Musculoskeletal comorbidity	2.45	1.64–3.67	0.000	2.75	1.83–4.12	0.000	2.12	1.43–3.14	0.000
Referred pain to the leg	1.98	1.44–2.73	0.000	2.31	1.67–3.19	0.000	1.62	1.18–2.24	0.003
Spinal morning stiffness (>30 min)	3.71	2.48–5.54	0.000	4.65	2.98–7.27	0.000	2.54	1.73–3.73	0.000
Pain during spinal rotation	2.53	1.80–3.56	0.000	2.32	1.65–3.27	0.000	2.03	1.44–2.86	0.000
Patients' expectation to recover	0.32	0.23–0.45	0.000	0.35	0.25–0.48	0.000	0.25	0.18–0.36	0.000
Depressive symptomatology (CES-D 0-60)	1.07	1.05–1.10	0.000	1.11	1.08–1.14	0.000	1.05	1.02–1.07	0.000
Kinesiophobia (FABQ-PA 0-28)	1.08	1.05–1.12	0.000	1.11	1.08–1.14	0.000	1.07	1.04–1.10	0.000
Pain catastrophizing (PCS 0-52)	1.04	1.02–1.06	0.000	1.07	1.05–1.09	0.000	1.03	1.02–1.05	0.000

\* Pooled estimates from the univariable regression analysis in the imputed data sets.

**Table 3****Multivariable clinical prediction models for nonrecovery outcomes in older adults with back pain (Back Complaints in the Elders cohort, n = 675, imputed data sets).**

Predictors	Coefficient (SE)	R <sup>2</sup>	Optimism-adjusted R <sup>2</sup>	AUC (95% CI)	Optimism-adjusted AUC
<b>Persistent back pain model*</b>					
Constant	−4.47 (1.00)	0.35	0.32	0.80 (0.76-0.83)	0.79
Age (y)	0.03 (0.01)				
Chronic duration (>3 mo)	0.77 (0.24)				
Average back pain intensity in the past wk (NRS 0-10) + spline	0.24 (0.12)				
	−0.11 (0.11)				
Disability (RMDQ 0-24)	0.07 (0.02)				
Recent episode of back pain (previous 6 mo)	0.75 (0.24)				
Spinal morning stiffness (>30 min)	0.41 (0.25)				
Pain during spinal rotation	0.44 (0.22)				
Patients' expectation to recover	−0.83 (0.21)				
<b>Persistent disability model†</b>					
Constant	−6.34 (1.14)	0.49	0.46	0.86 (0.83-0.89)	0.85
Age (y)	0.04 (0.01)				
Chronic duration (>3 mo)	0.67 (0.29)				
Disability (RMDQ 0-24) + spline	0.34 (0.06)				
	−0.18 (0.07)				
Recent episode of back pain (previous 6 mo)	0.43 (0.26)				
Musculoskeletal comorbidity	0.55 (0.26)				
Spinal morning stiffness (>30 min)	0.55 (0.28)				
Patients' expectation to recover	−0.88 (0.23)				
Depressive symptomatology (CES-D 0-60) + spline	0.09 (0.04)				
	−0.08 (0.06)				
Pain catastrophizing (PCS 0-52)	0.02 (0.01)				
<b>Perceived nonrecovery model‡</b>					
Constant	−3.60 (0.95)	0.32	0.30	0.78 (0.74-0.82)	0.77
Age (y)	0.03 (0.01)				
Chronic duration (>3 mo)	0.62 (0.25)				
Disability (RMDQ 0-24) + spline	0.13 (0.05)				
	−0.07 (0.05)				
Recent episode of back pain (previous 6 mo)	1.16 (0.23)				
Pain during spinal rotation	0.38 (0.21)				
Patients' expectation to recover	−0.99 (0.20)				

\* Penalized persistent back pain model (with adjusted coefficients; slope 0.94) is  $-4.23 + 0.03 \times \text{age} + 0.73 \times \text{chronic duration} + 0.23 \times \text{back pain intensity} - 0.10 \times \text{back pain intensity spline} + 0.07 \times \text{disability} + 0.71 \times \text{recent episode} + 0.38 \times \text{spinal morning stiffness} + 0.42 \times \text{pain during spinal rotation} - 0.79 \times \text{patients' expectation to recover}$ .

† Penalized persistent disability model (with adjusted coefficients; slope 0.94) is  $-5.97 + 0.04 \times \text{age} + 0.63 \times \text{chronic duration} + 0.32 \times \text{disability} - 0.17 \times \text{disability spline} + 0.40 \times \text{recent episode} + 0.52 \times \text{musculoskeletal comorbidity} + 0.52 \times \text{spinal morning stiffness} - 0.83 \times \text{patients' expectation to recover} + 0.08 \times \text{depressive symptomatology} - 0.08 \times \text{depressive symptomatology spline} + 0.02 \times \text{pain catastrophizing}$ .

‡ Penalized perceived nonrecovery model (with adjusted coefficients; slope 0.96) is  $-3.46 + 0.03 \times \text{age} + 0.60 \times \text{chronic duration} + 0.13 \times \text{disability} - 0.07 \times \text{disability spline} + 1.11 \times \text{recent episode} + 0.36 \times \text{pain during spinal rotation} - 0.95 \times \text{patients' expectation to recover}$ .

expectation to recover from their back pain. The persistent back pain model furthermore included pain intensity, spinal morning stiffness, and pain during spinal rotation as predictors. The persistent disability model additionally contained 2 physical factors, spinal morning stiffness and musculoskeletal comorbidity, as well as 2 psychosocial factors, namely, depressive symptomatology and pain catastrophizing. The perceived nonrecovery model included 1 additional predictor: pain during spinal rotation.

All 3 developed models exhibited good performance (R<sup>2</sup> range: 0.32-0.49; AUC range: 0.78-0.86) likewise the optimism-adjusted models (optimism-adjusted R<sup>2</sup> range: 0.30-0.46, respectively, AUC range: 0.77-0.85; **Table 3**). The Hosmer–Lemeshow was nonsignificant in all models, which means calibration was adequate; calibration plots are presented in **Figure 1A–C**. The best performing model is the one predicting persistent disability, which showed good discrimination (AUC 0.86, 95% CI 0.83-0.89) and explained variance (R<sup>2</sup> 49%) also after internal validation (optimism-adjusted R<sup>2</sup> 46% and AUC 0.85). The other 2 models also performed well after internal validation (**Table 3**).

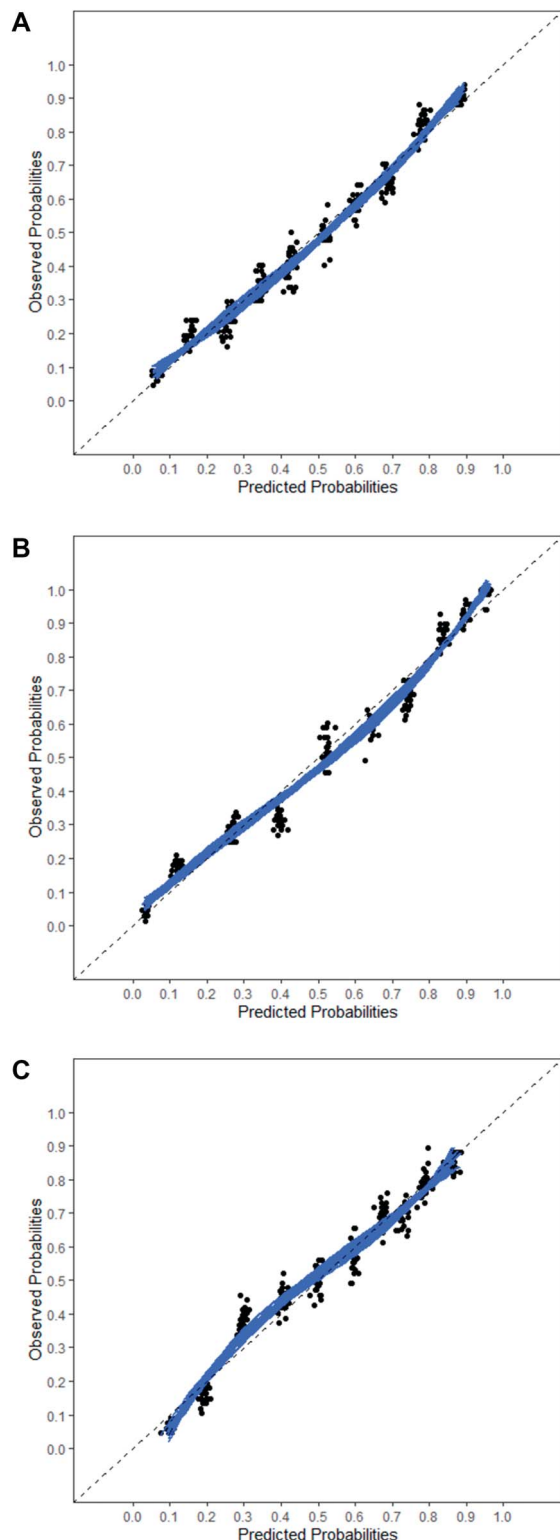
### 3.4. Sensitivity analyses

Results for the complete case analysis are presented in **Table 4**. The 3 models showed similar performance measures (R<sup>2</sup> range: 0.31-0.47; AUC range: 0.78-0.85) as the final multivariable prediction models in the imputed data sets, but some differences in predictors included in the models were identified. For the model-predicting persistent pain, musculoskeletal comorbidity was part of the model instead of spinal morning stiffness. The model-predicting persistent disability included pain during spinal rotation as an additional predictor; a recent episode of back pain and pain catastrophizing were no longer part of the model. The model-predicting perceived nonrecovery included all predictors of the imputed data set model, except for pain intensity.

## 4. Discussion

### 4.1. Statement of principle findings

An increasing age, chronic duration of symptoms, higher pain intensity and higher back pain–related disability, a recent back pain episode, musculoskeletal comorbidity, spinal morning



**Figure 1.** Calibration plots: (A) persistent back pain, (B) persistent disability, (C) perceived nonrecovery.

stiffness of over 30 minutes, pain during spinal rotation, a lower expectation of patients to recover from their back pain, depressive symptomatology, and pain catastrophizing were predictors in multivariable models in older adults with back pain for persistent nonrecovery after 6 and 12 months. Predictors consistently reported in all 3 models were age, chronic duration, disability, a recent back pain episode, and patients' recovery

expectations. Two of the 3 models also included spinal morning stiffness and pain during spinal rotation as predictors in the final prediction model. All models exhibited good explained variance, calibration, and discrimination after internal validation, with the disability model being the best performing one (Table 3). Sensitivity analysis with complete case data set showed some small differences in the predictors included in the models (Table 4).

#### 4.2. Comparison with the existing literature

There is a considerable amount of prognostic research in back pain, but study quality varies, and there is a lack of standardization in methods and outcomes.<sup>31</sup> Some of the existing prognostic models in back pain lack validation.<sup>32</sup> There are few prognostic models focusing on longer-term outcomes of back pain in the literature, and none of them with a specific focus on older adults, a population that is often excluded from participation in studies.<sup>50</sup> Previous studies in older adults examined predictor variables in a univariable analysis<sup>47</sup> or studied prognostic factors for nonrecovery after 3 months.<sup>49</sup> Therefore, it is not possible to compare it with other models in a similar population. It is also different from existing risk stratification tools, such as the STarT Back Tool,<sup>56</sup> that focus on stratified care by combining a treatment to a risk of outcome, whereas the current study solely focuses on prognosis.

Generally speaking, our findings are supported by the existing literature on prognostic factors for back pain in adults as well as existing clinical prediction models, and seem applicable to adults of all ages, such as having an older age at baseline<sup>19,31</sup>; a higher pain intensity,<sup>2,18,19,31</sup> higher levels of disability,<sup>2,8,19,31</sup> more previous episodes of back pain,<sup>2,10,18,19,31</sup> and a longer duration of back pain symptoms.<sup>2,10,18</sup> Interestingly, increasing age was a predictor in all 3 models. Dionne et al.<sup>12</sup> suggested in their review that the frequency of severe back pain increases with increasing age, which might be a reason for poor recovery from their back pain as well. Increasing age could then be seen as a mediator for often more severe and disabling back pain.

New findings are the importance of predictors such as spinal morning stiffness and pain during spinal rotation, 2 clinical features of back pain that were rather unknown as predictors previously. One hypothesis could be that with increasing age, these factors become more important for the persistence of pain and disability. Another reason could be that the development of symptomatic spinal osteoarthritis becomes apparent in a part of the older patients with back pain, and this disorder is a prognostic factor for poor recovery.<sup>59</sup>

Future research could establish whether this model is also valid in other similar populations and can be applied more widely in different countries and, perhaps, to different age groups. Once these models are externally validated, it will be possible to develop an online calculator for primary care clinicians to estimate the risk of poor prognosis. This could follow the example of the PICKUP tool for younger patients with acute low back pain, used to estimate their risk of chronic low back pain,<sup>57</sup> and may provide a more accurate tool for older adults, such as the patients included in our cohort.

#### 4.3. Strengths and limitations

This study is conducted in a real life primary care setting, and the results should be applicable in daily practice. The prediction models developed in this study had sufficient power to handle the number of

Table 4

**Multivariable clinical prediction models for nonrecovery outcomes in older adults with back pain (Back Complaints in the Elders cohort, n = 428, persistent back pain; n = 420, persistent disability; n = 433, perceived nonrecovery, complete case analysis).**

Predictors	Coefficient (SE)	R <sup>2</sup>	Optimism-adjusted R <sup>2</sup>	AUC (95% CI)	Optimism-adjusted AUC
<b>Persistent back pain model*</b>					
Constant	−5.02 (1.22)	0.31	0.27	0.78 (0.74-0.83)	0.76
Age (y)	0.03 (0.02)				
Chronic duration (>3 mo)	0.84 (0.27)				
Average back pain intensity in the past wk (NRS 0-10) + spline	0.38 (0.15)				
	−0.22 (0.14)				
Disability (RMDQ 0-24)	0.07 (0.02)				
Recent episode of back pain (previous 6 mo)	0.45 (0.28)				
Musculoskeletal comorbidity	0.42 (0.28)				
Pain during spinal rotation	0.57 (0.24)				
Patients' expectation to recover	−0.64 (0.25)				
<b>Persistent disability model†</b>					
Constant	−6.21 (1.32)	0.47	0.43	0.85 (0.82-0.89)	0.84
Age (y)	0.03 (0.02)				
Chronic duration (>3 mo)	0.68 (0.30)				
Disability (RMDQ 0-24) + spline	0.40 (0.07)				
	−0.25 (0.07)				
Musculoskeletal comorbidity	0.57 (0.30)				
Spinal morning stiffness (>30 min)	0.72 (0.35)				
Pain during spinal rotation	0.62 (0.27)				
Patients' expectation to recover	−1.00 (0.27)				
Depressive symptomatology (CES-D 0-60) + spline	0.14 (0.06)				
	−0.12 (0.06)				
<b>Perceived nonrecovery model‡</b>					
Constant	−4.44 (1.21)	0.36	0.32	0.80 (0.76-0.84)	0.78
Age (y)	0.03 (0.02)				
Chronic duration (>3 mo)	0.52 (0.27)				
Average back pain intensity in the past wk (NRS 0-10) + spline	0.24 (0.14)				
	−0.14 (0.14)				
Disability (RMDQ 0-24) + spline	0.17 (0.06)				
	−0.11 (0.06)				
Recent episode of back pain (previous 6 mo)	1.16 (0.29)				
Pain during spinal rotation	0.54 (0.25)				
Patients' expectation to recover	−1.20 (0.25)				

\* Penalized persistent back pain model (with adjusted coefficients; slope 0.90) is  $-4.54 + 0.03 \times \text{age} + 0.76 \times \text{chronic duration} + 0.34 \times \text{back pain intensity} - 0.20 \times \text{back pain intensity spline} + 0.06 \times \text{disability} + 0.40 \times \text{recent episode} + 0.38 \times \text{musculoskeletal comorbidity} + 0.51 \times \text{pain during spinal rotation} - 0.58 \times \text{patients' expectation to recover}$ .

† Penalized persistent disability model (with adjusted coefficients; slope 0.90) is  $-5.60 + 0.03 \times \text{age} + 0.61 \times \text{chronic duration} + 0.36 \times \text{disability} - 0.23 \times \text{disability spline} + 0.65 \times \text{spinal morning stiffness} + 0.51 \times \text{musculoskeletal comorbidity} + 0.56 \times \text{pain during spinal rotation} + 0.13 \times \text{depressive symptomatology} - 0.11 \times \text{depressive symptomatology spline} - 0.91 \times \text{patients' expectation to recover}$ .

‡ Penalized perceived nonrecovery model (with adjusted coefficients; slope 0.90) is  $-4.01 + 0.02 \times \text{age} + 0.47 \times \text{chronic duration} + 0.22 \times \text{back pain intensity} - 0.12 \times \text{back pain intensity spline} + 0.16 \times \text{disability} - 0.10 \times \text{disability spline} + 1.05 \times \text{recent episode} + 0.49 \times \text{pain during spinal rotation} - 1.08 \times \text{patients' expectation to recover}$ .

potential predictors initially included, as recent guidance on sample size calculation was used.<sup>43</sup> A general limitation in the field of back pain is that not yet one definition of “recovery” exists. It is known that how and when patients call themselves recovered is very personal and is different for almost each individual.<sup>28</sup> We tried to overcome this limitation by using 3 definitions of recovery on different domains. However, from the results, it is clear that most key predictors appeared in every model, irrespective of the outcome definition. This implies these definitions are highly related. Another limitation is that the BACE cohort did not contain the measures needed for the definition of widespread pain, a well-known prognostic factor in both adults and older adults.<sup>2,46</sup> However, the presence of comorbid musculoskeletal symptoms/pain could at least partially link to widespread pain and may be used as a proxy in this case.

## 5. Conclusion and implications for research and practice

This study presented 3 multivariable clinical prediction models for nonrecovery in older adults presenting with a new episode of care for back pain in general practice. The models displayed

good overall performance, calibration, and discrimination irrespective of the definition of outcome that was used. The model predicting persistent disability performed best. Because the BACE cohort study had a pragmatic design and was conducted in general practice, these models have potential to help primary care clinicians to determine which of their older patients may or may not recover from a back pain episode. Early identification and close follow-up of the patients at risk for a poor outcome could aid timely referral to intensified treatment. However, these results should first be externally validated in a similar population before they can be implemented in clinical primary care settings. If replicated, a next step could be the translation towards an easy applicable clinical prediction rule for primary care clinicians. Finally, a clinical impact analysis should be conducted to evaluate the effects on the management and recovery of older adults with back pain in primary care.

## Conflict of interest statement

The authors have no conflicts of interest to disclose.

## Acknowledgements

This study was funded by the Department of General Practice, Erasmus University Medical Center, Rotterdam, the Coolsingel Foundation, Rotterdam, and partly funded by a program grant of the Dutch Arthritis Foundation; all based in the Netherlands. The authors thank all study participants, participating GPs, the project team, and research assistants.

Ethical approval: the study protocol (NL24829.078.08) was approved by the Medical Ethics Committee of Erasmus University Medical Center, Rotterdam, the Netherlands.

## Article history:

Received 16 May 2020

Received in revised form 2 November 2020

Accepted 4 November 2020

Available online 15 December 2020

## References

- [1] Ahern M, Dean CM, Dear BF, Willcock SM, Hush JM. The experiences and needs of people seeking primary care for low-back pain in Australia. *Pain Rep* 2019;4:e756.
- [2] Artus M, Campbell P, Mallen CD, Dunn KM, van der Windt DAW. Generic prognostic factors for musculoskeletal pain in primary care: a systematic review. *BMJ Open* 2017;7:e012901.
- [3] Artus M, van der Windt D, Jordan KP, Croft PR. The clinical course of low back pain: a meta-analysis comparing outcomes in randomised clinical trials (RCTs) and observational studies. *BMC Musculoskelet Disord* 2014; 15:68.
- [4] Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. *PAIN* 1996;65: 71–6.
- [5] Campbell P, Foster NE, Thomas E, Dunn KM. Prognostic indicators of low back pain in primary care: five-year prospective study. *J Pain* 2013;14: 873–83.
- [6] Chiarotto A, Boers M, Deyo RA, Buchbinder R, Corbin TP, Costa LOP, Foster NE, Grotle M, Koes BW, Kovacs FM, Lin CWC, Maher CG, Pearson AM, Peul WC, Schoene ML, Turk DC, van Tulder MW, Terwee CB, Ostelo RW. Core outcome measurement instruments for clinical trials in nonspecific low back pain. *PAIN* 2018;159:481–95.
- [7] Chiarotto A, Deyo RA, Terwee CB, Boers M, Buchbinder R, Corbin TP, Costa LO, Foster NE, Grotle M, Koes BW, Kovacs FM, Lin CW, Maher CG, Pearson AM, Peul WC, Schoene ML, Turk DC, van Tulder MW, Ostelo RW. Core outcome domains for clinical trials in non-specific low back pain. *Eur Spine J* 2015;24:1127–42.
- [8] Chou R, Shekelle P. Will this patient develop persistent disabling low back pain? *JAMA* 2010;303:1295–302.
- [9] Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015;350:g7594.
- [10] da Silva T, Macaskill P, Mills K, Maher C, Williams C, Lin C, Hancock MJ. Predicting recovery in patients with acute low back pain: a Clinical Prediction Model. *Eur J Pain* 2017;21:716–26.
- [11] Deyo RA, Bryan M, Comstock BA, Turner JA, Heagerty P, Friedly J, Avins AL, Nedeljkovic SS, Nerenz DR, Jarvik JG. Trajectories of symptoms and function in older adults with low back disorders. *Spine* 2015;40:1352–62.
- [12] Dionne CE, Dunn KM, Croft PR. Does back pain prevalence really decrease with increasing age? A systematic review. *Age Ageing* 2006;35: 229–34.
- [13] Disease GBD, Injury I, Prevalence C. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018;392: 1789–858.
- [14] Dunn KM, Jordan KP, Croft PR. Contributions of prognostic factors for poor outcome in primary care low back pain patients. *Eur J Pain* 2011;15: 313–19.
- [15] Enthoven WTM, Koes BW, Bierma-Zeinstra SMA, Bueving HJ, Bohnen AM, Peul WC, van Tulder MW, Berger MY, Luijsterburg PAJ. Defining trajectories in older adults with back pain presenting in general practice. *Age Ageing* 2016;45:878–83.
- [16] Foster NE, Thomas E, Bishop A, Dunn KM, Main CJ. Distinctiveness of psychological obstacles to recovery in low back pain patients in primary care. *PAIN* 2010;148:398–406.
- [17] Grotle M, Foster NE, Dunn KM, Croft P. Are prognostic indicators for poor outcome different for acute and chronic low back pain consulters in primary care? *PAIN* 2010;151:790–7.
- [18] Hancock MJ, Maher CG, Latimer J, Herbert RD, McAuley JH. Can rate of recovery be predicted in patients with acute low back pain? Development of a clinical prediction rule. *Eur J Pain* 2009;13:51–5.
- [19] Hayden JA, Chou R, Hogg-Johnson S, Bombardier C. Systematic reviews of low back pain prognosis had variable methods and results: guidance for future prognosis reviews. *J Clin Epidemiol* 2009;62:781–96 e781.
- [20] Henschke N, Maher CG, Refshauge KM, Herbert RD, Cumming RG, Bleasel J, York J, Das A, McAuley JH. Prognosis in patients with recent onset low back pain in Australian primary care: inception cohort study. *BMJ* 2008;337:a171.
- [21] Hestbaek L, Leboeuf-Yde C, Manniche C. Low back pain: what is the long-term course? A review of studies of general patient populations. *Eur Spine J* 2003;12:149–65.
- [22] Heymans MW. Prediction Model Selection and Performance Evaluation in Multiple Imputed Datasets. *psfmi* 0.5.0, 2020. Available at: <https://mweymans.github.io/psfmi/>. Accessed October 1, 2020.
- [23] Heymans MW, Eekhout I. Applied Missing Data Analysis with SPSS and (R)Studio. Amsterdam, 2019. Available at: <https://bookdown.org/mweymans/bookmi/>. Accessed June 6, 2019.
- [24] Hoy D, Brooks P, Blyth F, Buchbinder R. The Epidemiology of low back pain. *Best Pract Res Clin Rheumatol* 2010;24:769–81.
- [25] Hoy D, March L, Brooks P, Woolf A, Blyth F, Vos T, Buchbinder R. Measuring the global burden of low back pain. *Best Pract Res Clin Rheumatol* 2010;24:155–65.
- [26] Hoy DG, Bain C, Williams G, March L, Brooks P, Blyth F, Woolf A, Vos T, Buchbinder R. A systematic review of the global prevalence of low back pain. *Arthritis Rheum* 2012;64:2028–37.
- [27] Hurwitz EL, Randhawa K, Yu H, Côté P, Haldeman S. The Global Spine Care Initiative: a summary of the global burden of low back and neck pain studies. *Eur Spine J* 2018;27:S796–801.
- [28] Hush JM, Refshauge K, Sullivan G, De Souza L, Maher CG, McAuley JH. Recovery: what does this mean to patients with low back pain? *Arthritis Rheum* 2009;61:124–31.
- [29] Kamper SJ, Stanton TR, Williams CM, Maher CG, Hush JM. How is recovery from low back pain measured? A systematic review of the literature. *Eur Spine J* 2011;20:9–18.
- [30] Kent P, Lauridsen HH. Managing missing scores on the Roland Morris disability questionnaire. *Spine* 2011;36:1878–84.
- [31] Kent PM, Keating JL. Can we predict poor recovery from recent-onset nonspecific low back pain? A systematic review. *Man Ther* 2008;13: 12–28.
- [32] McIntosh G, Steenstra I, Hogg-Johnson S, Carter T, Hall H. Lack of prognostic model validation in low back pain prediction studies: a systematic review. *Clin J Pain* 2018;34:748–54.
- [33] McRae M, Hancock MJ. Adults attending private physiotherapy practices seek diagnosis, pain relief, improved function, education and prevention: a survey. *J Physiother* 2017;63:250–6.
- [34] Menezes Costa LdC, Maher CG, Hancock MJ, McAuley JH, Herbert RD, Costa LOP. The prognosis of acute and persistent low-back pain: a meta-analysis. *CMAJ* 2012;184:E613–24.
- [35] Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. The TRIPOD statement: explanation and elaboration. *Ann Intern Med* 2015;162: W1–W73.
- [36] Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J, Mallett S. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *PROBAST: explanation and elaboration. Ann Intern Med* 2019;170: W1–W33.
- [37] Oliveira CB, Maher CG, Pinto RZ, Traeger AC, Lin C-WC, Chenot J-F, van Tulder M, Koes BW. Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. *Eur Spine Journal* 2018;27:2791–2803.
- [38] Pengel LHM, Herbert RD, Maher CG, Refshauge KM. Acute low back pain: systematic review of its prognosis. *BMJ* 2003;327:323.
- [39] 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.



- [40] Pincus T, Santos R, Breen A, Burton AK, Underwood M. A review and proposal for a core set of factors for prospective cohorts in low back pain: a consensus statement. *Arthritis Rheum* 2008;59:14–24.
- [41] Pincus T, Vogel S, Burton AK, Santos R, Field AP. Fear avoidance and prognosis in back pain: a systematic review and synthesis of current evidence. *Arthritis Rheum* 2006;54:3999–4010.
- [42] Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977;1:385–401.
- [43] Riley RD, Snell KI, Ensor J, Burke DL, Harrell FE, Jr, Moons KG, Collins GS. Minimum sample size for developing a multivariable prediction model: PART II—binary and time-to-event outcomes. *Stat Med* 2019;38:1276–96.
- [44] Riley RD, van der Windt DA, Croft P, Moons KG. Prognosis research in health care, concepts, methods, and impact. Oxford: Oxford University Press, 2019.
- [45] 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;8:141–4.
- [46] Rundell SD, Patel KV, Krook MA, Heagerty PJ, Suri P, Friedly JL, Turner JA, Deyo RA, Bauer Z, Nerenz DR, Avins AL, Nedeljkovic SS, Jarvik JG. Multisite pain is associated with long-term patient-reported outcomes in older adults with persistent back pain. *Pain Med* 2019;20:1898–1906.
- [47] Rundell SD, Sherman KJ, Heagerty PJ, Mock CN, Dettori NJ, Comstock BA, Avins AL, Nedeljkovic SS, Nerenz DR, Jarvik JG. Predictors of persistent disability and back pain in older adults with a new episode of care for back pain. *Pain Med* 2017;18:1049–62.
- [48] Rundell Sean D, Sherman Karen J, Heagerty Patrick J, Mock Charles N, Jarvik Jeffrey G. The clinical course of pain and function in older adults with a new primary care visit for back pain. *J Am Geriatr Soc* 2015;63:524–30.
- [49] Scheele J, Enthoven WT, Bierma-Zeinstra SM, Peul WC, van Tulder MW, Bohnen AM, Berger MY, Koes BW, Luijsterburg PA. Course and prognosis of older back pain patients in general practice: a prospective cohort study. *PAIN* 2013;154:951–57.
- [50] Scheele J, Enthoven WTM, Bierma-Zeinstra SMA, Peul WC, Tulder MW, Bohnen AM, Berger MY, Koes BW, Luijsterburg PAJ. Characteristics of older patients with back pain in general practice: BACE cohort study. *Eur J Pain* 2014;18:279–87.
- [51] Scheele J, Luijsterburg PA, Bierma-Zeinstra SM, Koes BW. Course of back complaints in older adults: a systematic literature review. *Eur J Phys Rehabil Med* 2012;48:37986.
- [52] Scheele J, Luijsterburg PA, Ferreira ML, Maher CG, Pereira L, Peul WC, van Tulder MW, Bohnen AM, Berger MY, Bierma-Zeinstra SM, Koes BW. Back complaints in the elders (BACE): design of cohort studies in primary care: an international consortium. *BMC Musculoskelet Disord* 2011;12:193.
- [53] Steyerberg EW, Moons KGM, van der Windt DA, Hayden JA, Perel P, Schroter S, Riley RD, Hemingway H, Altman DG. For the PG. Prognosis research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med* 2013;10:e1001381.
- [54] Sullivan MJL, Bishop SR. The pain catastrophizing scale: development and validation. *Psychol Assess* 1995;7:524–32.
- [55] Suri P, Delaney K, Rundell SD, Cherkin DC. Predictive validity of the STaRT back tool for risk of persistent disabling back pain in a U.S. Primary care setting. *Arch Phys Med Rehabil* 2018;99:1533–9.e1532.
- [56] Traeger A, McAuley JH. STaRT back screening tool. *J Physiother* 2013;59:131.
- [57] Traeger AC, Henschke N, Hubscher M, Williams CM, Kamper SJ, Maher CG, Moseley GL, McAuley JH. Estimating the risk of chronic pain: development and validation of a prognostic model (PICKUP) for patients with acute low back pain. *PLoS Med* 2016;13:e1002019.
- [58] Turner JA, Shortreed SM, Saunders KW, LeResche L, Berlin JA, Korff MV. Optimizing prediction of back pain outcomes. *PAIN* 2013;154:1391–401.
- [59] van den Berg R, Chiarotto A, Enthoven WT, de Schepper EI, Oei EHG, Koes B, Bierma-Zeinstra SM. Clinical and radiographic features of spinal osteoarthritis predict long-term persistence and severity of back pain in older adults. *Ann Phys Rehabil Med* 2020; S1877-0657(20)30153-6.
- [60] van der Gaag WH, Enthoven WTM, Luijsterburg PAJ, van Rijkevorsel-Scheele J, Bierma-Zeinstra SMA, Bohnen AM, van Tulder MW, Koes BW. Natural history of back pain in older adults over five years. *J Am Board Fam Med* 2019;32:781–9.
- [61] Von Korff M, Jensen MP, Karoly P. Assessing global pain severity by self-report in clinical and health services research. *Spine* 2000;25:3140–51.
- [62] 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.
- [63] Ware JE, Jr, Gandek B. Overview of the SF-36 health survey and the international quality of life assessment (IQOLA) project. *J Clin Epidemiol* 1998;51:903–12.
- [64] Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J, Mallett S, Groupdagger P. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019;170:51–8.
- [65] Wong AY, Lauridsen HH, Samartzis D, Macedo L, Ferreira PH, Ferreira ML. Global consensus from clinicians regarding low back pain outcome indicators for older adults: pairwise wiki survey using crowdsourcing. *JMIR Rehabil Assist Technol* 2019;6:e11127.