



# Development of a standard set of outcome measures for non-specific low back pain in Dutch primary care physiotherapy practices: a Delphi study

A. C. Verburg<sup>1</sup> · S. A. van Dulmen<sup>1</sup> · H. Kiers<sup>2,3</sup> · M. W. G. Nijhuis-van der Sanden<sup>1</sup> · P. J. van der Wees<sup>1</sup>

Received: 14 December 2018 / Revised: 22 February 2019 / Accepted: 24 March 2019 / Published online: 19 April 2019  
© The Author(s) 2019

## Abstract

**Purpose** To develop a clinical standard set of outcome measures that are accepted for relevance and feasibility by stakeholders and useful for (a) interaction between patient and the professional, e.g. shared decision-making in goal-setting, monitoring and feedback based on outcomes, (b) internal quality improvement, and (c) external transparency in patients with non-specific low back pain (NSLBP) in primary care physical therapy.

**Methods** We used a consensus-driven modified RAND-UCLA Delphi method in seven steps with panellists (patients, representatives of patient and physiotherapy associations, researchers, policy makers, health insurers): (1) literature search, (2) first online survey, (3) patient interviews, (4) an experts meeting, (5) a consensus meeting, (6) second online survey, and (7) final approval of an advisory board. Steps 1–4 resulted in potential outcome measures. In the consensus meeting after discussion panellists voted for inclusion per measure. In the second online survey the final standard set was rated on relevance and feasibility on a 9-point Likert scale; when the median score was  $\geq 7$ , the standard set was accepted and finally approved.

**Results** Thirteen draft outcome measures were rated and discussed, and finally, six outcome measures were accepted. The standard set includes the Quebec Back Pain Disability Scale, Oswestry Disability Index, Patient-Specific Functional Scale, Numeric Pain Rating Scale, Global Perceived Effect (GPE-DV), and the STarT Back Screening Tool (SBT).

**Conclusion** This study presents a standard set of outcome measures for patients with NSLBP in primary care physiotherapy accepted for relevance and feasibility by stakeholders. The standard set is currently used in daily practice and tested on validity and reliability in a pilot study.

## Graphical abstract

These slides can be retrieved under Electronic Supplementary Material.

**Spine Journal**

### Key points

Key words: nonspecific low back pain, Outcome measures, Patient reported outcome measures, Patient reported experience measures, Delphi procedure

1. Target group: Patients with nonspecific low back pain
2. Focus of the standard set: Outcome measures
3. Perceived outcomes: PROMs
4. Patient experiences: PREMs
5. Methods: Delphi

Verburg AC, van Dulmen SA, Kiers H, Nijhuis-van der Sanden MWG, van der Wees PJ (2019) Development of a standard set of outcome measures for nonspecific low back pain in Dutch primary care physiotherapy practices: a Delphi study. *Eur Spine J*. Springer

**Spine Journal**

### Final outcome set for patients with NSLBP

Measurement	Low risk profile*		Medium/high risk profile*		
	Intake	End of treatment	Intake	Every six weeks	End of treatment
Quebec Back Pain Disability Scale			X	X	X
Oswestry Disability Index (sub-pilot)			X	X	X
Patient Specific Functioning Scale	X	X	X	X	X
Numeric Pain Rating Scale	X	X	X	X	X
Global Perceived Effect Dutch Version		X			X

\*The Subgroups for Targeted Treatment (STarT) Back Screening Tool is used for allocating patients in low, medium or high risk

Verburg AC, van Dulmen SA, Kiers H, Nijhuis-van der Sanden MWG, van der Wees PJ (2019) Development of a standard set of outcome measures for nonspecific low back pain in Dutch primary care physiotherapy practices: a Delphi study. *Eur Spine J*. Springer

**Spine Journal**

### Take Home Messages

1. This study presents a consensus based clinical standard set of outcome measures that is accepted for relevance and feasibility by stakeholders.
2. The outcome set is currently used in daily practice and tested on validity and reliability.
3. All stakeholders stay engaged during further validation and feasibility of the accepted standard set.

Verburg AC, van Dulmen SA, Kiers H, Nijhuis-van der Sanden MWG, van der Wees PJ (2019) Development of a standard set of outcome measures for nonspecific low back pain in Dutch primary care physiotherapy practices: a Delphi study. *Eur Spine J*. Springer

**Keywords** Non-specific low back pain · Outcome measures · PROMs · PREMs · Delphi

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00586-019-05962-x>) contains supplementary material, which is available to authorized users.

Extended author information available on the last page of the article

## Introduction

Quality of care has been defined as: *doing the right thing, at the right time, in the right way, for the right person, and having the best possible results* [1]. The challenge is how to measure quality of care in daily practice. Results at the level of the patient's health status can be measured with patient-reported outcome measures (PROMs), patient-reported experience measures (PREMs), and/or physical performance measures. PROMs are questionnaires or single-item scales measuring aspects of a patient's health status directly reported by the patient, e.g. perceived pain. PREMs are questionnaires measuring the experience of patients with health care, e.g. communication with the health care professional. Physical performance measures are clinical tests to measure physical function, e.g. 6-min walking test. Outcome measures should be well developed and unidimensional in order to generate information regarding the construct of interest [2]. Using such combined outcomes in daily practice is proposed to facilitate the interaction between patient and health care professional, including the process of shared decision-making, goal-setting, and evaluation of treatment effects [3, 4].

Outcomes measurement can also be useful to provide transparency about the process and the intervention effect at the group level in order to facilitate quality improvement trajectories, to provide information for patients, and for pay-for-performance purposes [5–7]. For successful implementation of outcome measures, patients and health care professionals need to accept a common set of outcomes to be measured as having added value in daily practice. In physiotherapy practice, multiple outcome measures are being used in clinical decision-making. Routine data collection in daily practice opens the opportunity for establishing large data sets with patient outcomes. Standardization of these measurements is necessary to enable comparison of intervention effects [3, 8].

Currently, limited data are available about the quality of daily care of patients with non-specific low back pain (NSLBP) treated in physiotherapy practices in the Netherlands. Therefore, this study focuses on the development of a standard set of outcome measures for NSLBP, the most common health condition of patients visiting physiotherapy in primary care practice. The final set of measures should be accepted as having added value in clinical practice and will have to be useful to compare the outcomes at the level of the individual patient, and for measuring and improving quality of physiotherapists and their practices.

Previous international studies showed several initiatives for developing outcome sets for low back pain [9–14]. Most of these core outcome sets were developed for clinical trial purposes and have not been tested with regard to relevance and feasibility in the evaluation of quality of care in daily

practice. Few studies on NSLBP in physiotherapy showed a good relationship of higher guideline adherence with better outcomes and less utilization of care [3, 15]. This stresses the value of gaining insight into outcomes on a larger scale. Successful implementation of outcome measures in daily practice can be improved by stakeholder engagement in quality improvement initiatives [16, 17]. It is therefore important to include all relevant stakeholders in Dutch physiotherapy concerning NSLBP in the development of the current standard set of outcomes.

In this study, NSLBP was defined as pain and discomfort, localized below the costal margin and above the inferior gluteal folds, with or without leg pain, and not caused by specific pathology [18]. NSLBP is an example of a patient group with high variation in level of recovery and time to recover varying from one day to multiple years. There is an increasing popularity for stratification of patients with NSLBP in subgroups, taking into account differences in characteristics based on prognostic profiles [19–22]. Outcome sets combined with stratified care will be more precise and useful for the development of quality indicators and better accepted for quality improvement [23].

The aim of this study was to develop a clinical standard set of outcome measures in patients with NSLBP—taking into account classification in clinically relevant subgroups—that is accepted for relevance and feasibility by stakeholders, and deemed useful for (a) interaction between patient and health care professional, e.g. shared decision-making in goal-setting and monitoring and feedback based on outcomes, (b) internal quality improvement, and (c) external transparency of primary care physical therapist practices.

## Methods

### Design

This study used a mixed method design in Dutch physiotherapy practices. The study was conducted between October 2016 and July 2017. An advisory board was formed with representatives of the Dutch Patient Association for Back Pain (NVVR), the Royal Dutch Society for Physiotherapy (KNGF), the Association for Quality in Physiotherapy (SKF) and two representatives of health insurance companies in the Netherlands (CZ & DFZ) to monitor and evaluate the process and to facilitate the implementation by providing and receiving information from stakeholder groups.

We used a consensus-driven modified RAND-UCLA Delphi method to select relevant outcomes [24] in seven separate steps (see Table 1). Informed consent was obtained from all individual participants included in the study, and all procedures were conducted according to the Declaration of Helsinki.

**Table 1** Steps during the consensus-driven modified Delphi method

Step	Participants	Goal	Result
1. Explorative review of the literature	Research team (JL, AV, SD, RN, and PW)	Identification of valid structure, process, and outcome (PROMs, PREMs, and physical performance measures) measures for NSLBP that are described in scientific literature Screening for profiles in subgroups related to the course of recovery based on prognostic factors	Selection and analysis of all eligible structure, process, and outcome measures on their validity and reliability Selection of eligible prognostic profiles for NSLBP
2. First online survey round	32 out of 43 Dutch physiotherapists with ample treatment or scientific experience on NSLBP	Anonymous rating of structure, process and outcome measures with a 9-point Likert scale on relevance and feasibility	The median score of every measure was calculated
3. Expert committee	4 senior physiotherapists of the first online survey round	Results of the online survey round 1 were interpreted and discussed with experts	First interpretation of the results of the first online survey presented and discussed
4. Interview with patients	6 patients with NSLBP and treated by physiotherapists last year	Patients views of measurements in clinical practice	The patients' perspectives on the use of measurement instruments which were presented during the introduction of the consensus meeting (step 5)
5. Consensus meeting	16 out of 43 participants of the first online survey and representatives of patients, KNGF, SKF, the advisory board and policy makers	Nominal group technique was used to discuss the draft set. Finally, the participants voted on feasibility and relevance	Measures were included if 80% or higher voted yes for inclusion. Between 60 and 80% yeses were deferred and re-rated in the second online survey, and between 0 and 60% yeses were excluded
6. Second online survey round	29 out of 43 participants of the first online survey (step 2)	The participants rated the second draft set and if needed alternatives on a 9-point Likert scale	Measures were included when all participants rated a median of 7 or higher on the 9-point Likert scale
7. Final approval of the advisory board	Advisory board (MJS, LV, HW, AW, AT, and ML)	The advisory board was asked to accept the final set	Final acceptance of the set

### *Step 1. Explorative review of the literature*

We searched in the Guideline International Network (G-I-N) and PEDro database for outcome measures based on PROMs, PREMs and physical performance measures with adequate psychometric properties, including reliability, validity, and responsiveness [25, 26]. All multi- and mono-disciplinary Dutch and international clinical practice guidelines for physiotherapists, general practitioners and medical specialists were included. We also searched websites of organizations developing clinical practice guidelines (see ‘Appendix A’). Based on the identified guidelines, reference tracking was performed. We preferred outcome measures that were already used in daily practice in the Netherlands. We also searched for structure, process and outcome measures in existing indicator sets, see ‘Appendix B’ for the search string. We used a pragmatic explorative approach and did not aim at conducting a systematic review of the literature. In the next step we analysed all eligible measures on their validity and reliability. The following information was gathered: type of measure (process, structure, outcome), type of questionnaire/instrument, targeted patient group, content of the questionnaire/instrument, time to complete the questionnaire/instrument, the minimal clinically important difference (MCID), domain (e.g. pain), related measurements, whether the questionnaire/instrument was already translated in Dutch, and supporting literature.

We used the prognostic profiles of the KNGF guideline as primary classification for subgroups related to the course recovery based on prognostic factors [27]. Then, we used the same literature search as for the outcome measures to compare the prognostic profiles of the KNGF with multi- and mono-disciplinary Dutch and international clinical practice guidelines. If necessary, we combined useful elements of different guidelines with the profiles of the KNGF guideline. In the identified guidelines, reference tracking was performed. Additionally, the PubMed database was screened between January 2012 and December 2016 for systematic reviews about individual prognostic factors in NSLBP, see ‘Appendix B’ for the search string. The individual prognostic factors were used as addition on the search to prognostic profiles. After the screening we selected all useful factors for prognostic profiles to classify subgroups for NSLBP.

### *Step 2. First online survey round*

We recruited 43 Dutch physiotherapists via contact networks of the KNGF and SKF for participation in the Delphi rounds; this was a purposive sample. All participants needed to have ample treatment experience in patients with NSLBP, or experience in scientific research on NSLBP, or both. The goal of this step was to rate all in step 1 selected measures with a 9-point Likert scale on relevance and feasibility. Afterwards, the participants rated the appropriateness and feasibility of prognostic profiles for NSLBP. We conducted the online survey in LimeSurvey, version 2.06.

### *Step 3. Expert committee*

We invited four participants with complementary expertise of the online survey to join an expert committee to discuss the results of the online survey and rated the outcome set and prognostic profiles on its content validity and reliability.

### *Step 4. Patient interviews*

We invited six patients who were treated by physiotherapists for NSLBP in the past year. The patients were recruited via a convenience sample of six physiotherapy practices. Each physiotherapy practice included one adult patient that was treated for NSLBP in the last year. When the patient agreed on the informed consent, the physiotherapist gave the contact information to the researcher. Short semi-structured telephone interviews of approximately 30 min were held by two researchers (KV and JL), and a topic list was used. The aim of the interviews was to gain insight into the patient perspective on relevance and feasibility on the use of questionnaires (PROMs and PREMs), physical performance measures, process and structure measures, and to what extent measurements can be used for improving the quality of care. The interviews were audio-recorded, transcribed verbatim and analysed using thematic analysis. KV and JL independently analysed the interviews and assigned codes within and between the interviews. Afterwards, the assigned codes were compared and grouped together in greater categories and themes. The most important themes were presented during the consensus meeting.

### *Step 5. Consensus meeting*

All participants of the online survey were invited in a three-hour consensus meeting, together with policy makers and members of the advisory board. We used the nominal group technique (NGT) to structure the meeting [28, 29]. The NGT is specifically designed and widely used for consensus statements between experts in a certain topic [28, 29]. The different steps in the NGT helps to give all participants a voice in the consensus process [30]. During these steps the participants rated, discussed and then re-rated the eligible structure, process and outcome measures. We discussed all measures scored in the first online survey and presented the results of steps 2–4, followed by a second on-site rating of the relevance and feasibility of the measures. The measures were included in the standard set if the total votes scored 80% or higher on yes/no rating. All measures that were scored between 60% and 80% in the on-site rating were deferred for discussion in the second online survey. All measures that received between 0% and 60% of the votes were excluded.

### *Step 6. Second online survey round*

All participants were invited for the second online round. All outcome measures that received between 60% and 80% of the votes were re-rated on relevance, and if needed, alternatives that were discussed in step 5 were rated. Based on

**Table 2** Result of step 2: first online survey and step 5: consensus meeting

Measures	Type of measure	Relevance (step 2)	Feasibility (step 2)	Inclusion for the final standard set (step 5) <sup>b</sup>	
		Median	Median	Yes (%)	No (%)
Quebec Back Pain Disability Scale (QBPD)	Outcome	7	7	80 <sup>c</sup>	20
Oswestry Disability Index (ODI)	Outcome	5	7	64 <sup>b</sup>	36
Patient-Specific Functional Scale (PSFS)	Outcome	7	8	76 <sup>b</sup>	24
Numeric Pain Rating Scale (NPRS)	Outcome	7	9	100 <sup>c</sup>	0
Global Perceived Effect—Dutch Version (GPE-DV)	Outcome	7	8	88 <sup>c</sup>	12
Improvement of activities	Outcome	6	7	0	100
Patient-Reported Experience Measures (PREM)	Outcome	7	5	19	81
Subgroups for Targeted Treatment (STarT) Back Screening Tool (SBT)	Screening tool	4	7	93 <sup>c</sup>	7
Tampa Scale of Kinesiophobia (TSK)	Outcome	4	6	0	100
Number of treatment sessions	Outcome	5	8	Not rated <sup>a</sup>	Not rated <sup>a</sup>
History taking described in the EHR	Process	6	7	0	100
Treatment plan described in the EHR	Process	6	7	0	100
Prognostic profiles from guidelines	Prognostic profiles	7	7	Not rated <sup>a</sup>	

*EHR* Electronic Health Record

<sup>a</sup>The number of treatment sessions and prognostic profiles was not rated because the panellists suggested an alternative instrument

<sup>b</sup>Re-rated in the second online survey

<sup>c</sup>Final inclusion in the standard set after rating

these results, we developed a final outcome set. We concluded the set and the prognostic profiles as being accepted when all panellists rated a median of 7 or higher on a 9-point Likert scale.

#### *Step 7. Final approval of the advisory board*

Finally, the advisory board was asked to accept the final outcome set. The goal was to inform every stakeholder about the conclusions and implications of the study and to increase the acceptability. All representatives were asked to take the responsibility for communication of the results of the study in their own organization.

## Results

### Participants

In step 2, 32 of the 43 panellists (response rate: 70%) completed the survey. The mean age of the panellists was 42 years, and 75.8% were men. In step 3, four expert physiotherapists (FM, DH, BM, and BH) accepted to participate. During step 4, six semi-structured interviews with six patients with NSLBP were held. The age ranged from 42 to 73 years with an average age of 56; four of the participants were men. In step 5, the consensus meeting, 16 of the 43 physiotherapists and researchers participated (response rate: 37%), as well as seven policy makers and members

of the advisory board. The patient representatives were not able to join at the meeting. For the second online survey in step 6, 29/43 (response rate: 68%) respondents participated. Finally, in step 7, the five members of the advisory board participated.

#### *Step 1. Explorative review of the literature*

We identified 27 measures, of which 13 were eligible for further investigation: six PROMs, two PREMs, two additional outcome measures, two process measures, and one screening tool (see Table 2).

The reasons for exclusion of the 14 measures were as follows: not familiar in the Netherlands, not developed and useful for NSLBP, not primarily advised by guidelines, and not for physiotherapy primary care purposes. For a clear description of all included measures and an overview of all excluded measures, see 'Appendix C'.

#### *Development of prognostic profiles*

We identified 19 guidelines, of which ten described useful information about prognostic profiles for NSLBP or individual prognostic factors [21, 27, 31–39]. The remaining nine guidelines were focused on specific pathology. To develop prognostic profiles we focused primarily on the Dutch KNGF guideline and compared it with other guidelines [27]. The majority of the guidelines specified two or three prognostic patient profiles based on the expected time of recovery. Some guidelines did not provide prognostic profiles in a table, but described them narratively [35–37].

Furthermore, all guidelines described individual prognostic factors that are associated with the course of recovery. See ‘Appendix D’ for a summary of all useful components.

Based on the outcomes of the literature review, we distinguished three prognostic profiles for NSLBP (A, B, and C), all profiles containing four characteristics. These characteristics are generally based on prognostic (psychosocial) factors of the KNGF guideline. Some guidelines described the expected time of recovery in weeks; this was added to our prognostic profiles. We identified prognostic factors based on back pain-related factors, individual factors, work-related factors, and psychosocial factors. The profiles are described in ‘Appendix E’.

#### *Step 2. First online survey round*

Of the 13 outcome measures that were rated, five outcome measures scored a median of  $\geq 7$  on *relevance* and on *feasibility* nine outcome and process measures scored a median of  $\geq 7$ . The prognostic profiles scored a median of  $\geq 7$  on *relevance* and *feasibility*; see Table 2 for a more specified overview of the survey.

#### *Step 3. Expert committee*

During the expert meeting all 13 selected outcome and process measures were discussed for validity and reliability. The experts accepted the prognostic profiles as having added value in daily practice. They made some suggestions, as to their opinion prognostic profiles are still not accurate enough to predict individuals who develop chronic pain or not. For example, prognostic factors of a patient in profile A can also be seen in profile B. Acute, sub-acute, and chronic low back pain should not be used in the prognostic profiles. The experts suggested selecting outcome measures per profile. The experts stated that it will be necessary to perform a solid pilot study to test the selected outcome and process measures on feasibility before the outcome set can be considered valid and reliable for quality improvement purposes.

#### *Step 4. Patient interviews*

The following themes were identified: (1) patient satisfaction, (2) administration, (3) number of treatment sessions, (4) transparency, and (5) PROMs. Almost every patient agreed that satisfaction about the given treatment and treatment effect is relevant for quality evaluation purposes. Clinical record keeping is important to monitor the effect of treatment, to support a colleague during takeovers, and is valuable for evaluating quality of care, but should also be short and brief. The number of treatment sessions can be useful to evaluate quality of care, depending on the patient group. Some patients stated that transparency about outcomes of care could help them to choose health care professionals, and other patients said that they preferred the advice of a doctor, therapist, or family member. There were different opinions whether PROMs were relevant for quality evaluation purposes. Most patients stated that the readability of PROMs is good. Some patients said that pain

and functional problems were useful elements to score in PROMs but also psychosocial factors as this may have influence on the effect of treatment. See ‘Appendix F’ for the themes, categories, and codes.

#### *Step 5. Consensus meeting*

After presentation of the results of previous steps, the panellists could vote per measure with yes/no whether it should be added to the final outcome set; all results are presented in Table 2. The Quebec Back Pain Disability Scale (QBPDS), Numeric Pain Rating Scale (NPRS), Global Perceived Effect—Dutch Version (GPE-DV), and STarT Back Screening Tool (SBT) were chosen to include to the final outcome set directly. The Oswestry Disability Index (ODI), Patient-Specific Functional Scale (PSFS), number of treatment sessions, and prognostic profiles were scored between 60 and 80%, and the following suggestions were done during the consensus meeting for adapting the measures: the panellists suggested that the ODI is less common in the Netherlands compared to the QBPDS. Nevertheless, the ODI is widely used internationally and consists of good psychometric properties. The panellists suggested adding the ODI to the outcome set and compare the ODI with the QBPDS in a pilot study. The pilot study opens the opportunity to reflect on preferences of the field and to analyse the feasibility, acceptability, and responsiveness between both instruments. In the meantime, physiotherapists in the field can choose between the ODI and QBPDS. The panellists stated that the number of treatment sessions should be changed to measuring the total costs of the episode. The panellists suggested that the STarT Back Screening Tool (SBT) could replace the prognostic profiles.

#### *Step 6. Second online survey round*

The following four questions related to the remaining measures were rated on a 9-point Likert scale: ‘Do you agree that we use the SBT to classify patients in subgroups?’, ‘Do you agree that we add the PSFS in the outcome set?’, ‘Do you agree that the ODI will be tested in a pilot study in comparison with the QBPDS?’, and ‘Do you agree that treatment costs should be evaluated but will not be added in the outcome set?’. In the online survey we presented the final outcome set and prognostic profiles as in Table 3. All questions scored a median of 7 or higher, and therefore, the outcome set was accepted.

#### *Step 7. Final approval of the advisory board*

The final set with outcome measures was accepted by the advisory board. They accepted the outcome set by signing an official approval document.

**Table 3** Final set of measures accepted by all stakeholders

Measurement	Low-risk profile <sup>b</sup>		Medium-/high-risk profile <sup>b</sup>		
	Intake <sup>a</sup>	End of treatment <sup>a</sup>	Intake <sup>a</sup>	Every six weeks <sup>a</sup>	End of treatment <sup>a</sup>
Quebec Back Pain Disability Scale (QBPDS)			X	X	X
Oswestry Disability Index (ODI)(pilot)			X	X	X
Patient-Specific Functional Scale (PSFS)	X	X	X	X	X
Numeric Pain Rating Scale (NPRS)	X	X	X	X	X
Global Perceived Effect—Dutch Version (GPE-DV)		X			X

<sup>a</sup>Intake, every six weeks and end of treatment refer to the time that physiotherapists need to score the outcomes

<sup>b</sup>The Subgroups for Targeted Treatment (STarT) Back Screening Tool (SBT) is used for allocating patients in low-, medium-, or high-risk profile

## Discussion

This study presents a standard set of six clinical outcome measures in patients with NSLBP in primary care physiotherapy, which includes the Quebec Back Pain Disability Scale (QBPDS), Oswestry Disability Index (ODI), Patient-Specific Functional Scale (PSFS), Numeric Pain Rating Scale (NPRS), Global Perceived Effect—Dutch Version (GPE-DV), and the STarT Back Screening Tool (SBT). The outcome measures are aimed to be used for the interaction between patient and physiotherapist, for internal quality improvement, and for external transparency.

In our study, the STarT Back Screening Tool was selected to allocate patients in subgroups, while research showed that cautiousness is required with respect to interpretation of prognostic tools [20]. Karran et al. (2017) concluded that prognostic screening instruments in primary care scored poorly at assigning higher risk scores to individuals who develop chronic pain, than those who will not [20]. However, other researches showed that identifying subgroups of patients with NSLBP is still promising for future health care [40]. Multiple researchers support this vision [20–22]. During a pilot study, we should test whether the STarT Back Screening Tool is reliable and valid for classifying patients in subgroups.

The fundamental difference with existing outcome sets for low back pain is that this outcome set is accepted by stakeholders as having added value in daily care [9–14]. Therefore, this new standard set provides a more promising basis for the implementation of quality indicators in clinical practice. Stakeholder engagement is essential for successful implementation of quality improvement initiatives. In comparison with traditional Delphi methods we performed additional activities to reach consensus based on the RAND-UCLA appropriateness method. Along with the anonymous online surveys and consensus meeting, we conducted an expert meeting, interviewed patients, and

consulted an advisory board. With these steps, stakeholders were encouraged to use this outcome set in daily practice. In our study the patients were included to reflect on the selection of measures for the standard set and not as a separate qualitative study. The interviews were limited to six patients, and we did not reach data saturation, which could lead to bias. However, the interviews gained sufficient insights of the patients' views of measurements in clinical practice.

During the consensus study we found that the stakeholders and physiotherapists and others showed a positive attitude about developing the standard set and its described goals. This may not be representative for the total population of physiotherapists in the Netherlands. The panellists of this study may have been early adopters and open for quality improvement or external transparency. During implementation of this standard set we will need to anticipate that physiotherapists need more information about the benefits on standardization of outcome measurements to provide insight into and to compare intervention effects. Implementation strategies may need to be aimed at knowledge, skills, and attitudes of physiotherapists.

Due to pragmatic reasons we were not able to let panellists rate the outcome measures on a 9-point Likert scale during the consensus meeting, as preferred by the RAND-UCLA method [24]. The panellists voted with yes/no. Potentially, this may have influenced the voting and panellists could feel peer pressured with the used method. We verified whether the panellists felt comfortable about the procedure, and they agreed and felt safe to give their opinion. We do not expect that the results would have been different when using a Likert scale.

Before implementation of a standard set in daily practice, it is important to develop an infrastructure for collection of the data [3, 8]. For example, the standard set must be implemented in the Electronic Health Record (EHR) that physiotherapists use for their clinical record keeping [3]. This EHR must be connected to a secure central database before

the outcomes can be analysed. Also, the infrastructure must allow the possibility to give practices and physiotherapists feedback on outcomes and useful for quality improvement.

In this study we present a consensus-based standard set of outcome measures that is accepted for relevance and feasibility by stakeholders. Therefore, this standard outcome set provides a promising basis for further development of quality indicators in physiotherapy practice [41]. The standard set is currently used in daily practice and tested for validity and reliability in a pilot before it can be used for the development of quality indicators [7]. All stakeholders should stay engaged during further implementation of the standard set.

**Acknowledgements** We acknowledge research members: Joan Luites (JL) and Juliette Cruisberg (JC); expert committee: Francois Maissan (FM), David den Hartog (DH), Bob van den Meiracker (BM), and Marcel Heijmans (MH); advisory board: Annemarie Trompert (AT), Marije de Leur (ML), Leen Voogt (LV), Hans Wattel (HW), and Alida Wolters (AW). Abbreviations research members: Arie Cornelis Verburg (AV) Simone van Dulmen (SD) Ria Nijhuis-van der Sanden (RN) and Philip van der Wees (PW).

**Funding** Funding support for this study was provided by Dutch health insurer CZ.

## Compliance with ethical standards

**Conflict of interest** Author van der Wees received a research grant from Dutch health insurer CZ.

**Ethics approval** The protocol was approved by the Medical Ethical Committee of Radboud University Medical Centre (Registration # 2017-3154).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## Appendix A

Organization/institute	Website
The Royal Dutch Society for Physiotherapy (KNGF)	<a href="http://www.fysionet.nl">www.fysionet.nl</a>
Dutch Association for General Practitioners	<a href="http://www.nhg.nl">www.nhg.nl</a>
Stichting Miletus	<a href="http://www.stichtingmiletus.nl">www.stichtingmiletus.nl</a>
IQ Healthcare	<a href="http://www.IQhealthcare.nl">www.IQhealthcare.nl</a>
Dutch National Registry Physiotherapy	<a href="http://www.Landelijkedatabasefysiotherapie.nl">www.Landelijkedatabasefysiotherapie.nl</a>
Association for Quality in Physiotherapy (SKF)	<a href="http://www.keurmerkfysotherapie.nl">www.keurmerkfysotherapie.nl</a>
Dutch Guideline Database	<a href="http://www.Richtlijndatabase.nl">www.Richtlijndatabase.nl</a>

Organization/institute	Website
National Institute for Health and Care Excellence	<a href="http://www.nice.org.uk">www.nice.org.uk</a>
Scottish Intercollegiate Guidelines Network (SIGN)	<a href="http://www.sign.ac.uk">www.sign.ac.uk</a>
National Health Service (NHS)	<a href="http://www.nhs.uk">www.nhs.uk</a>
American Physical Therapy Association (APTA)	<a href="http://www.apta.org">www.apta.org</a>
Organisation for Economic Co-operation and Development (OECD)	<a href="http://www.oecd.org">www.oecd.org</a>
The International Consortium for Health Outcomes Measurement (ICHOM)	<a href="http://ichom.org">ichom.org</a>
Das Ärztliche Zentrum für Qualität in der Medizin (AZQ)	<a href="http://www.aezq.de">www.aezq.de</a>
Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen (aQua)	<a href="http://www.aqua-institut.de">www.aqua-institut.de</a>
National Quality Forum (NQF)	<a href="http://www.qualityforum.org">www.qualityforum.org</a>
National Health Service Scotland (NHSS)	<a href="http://www.indicators.scot.nhs.uk">www.indicators.scot.nhs.uk</a>
RAND Corporation	<a href="http://www.rand.org/">www.rand.org/</a>
Agency for Healthcare Research and Quality (AHRQ)	<a href="http://www.guidelines.gov">www.guidelines.gov</a>
Dutch Society for Cesar & Mensendieck Therapies	<a href="https://vvocm.nl/">https://vvocm.nl/</a>

## Appendix B

### Search for indicator sets

(low back pain[ti] AND (((Outcome Assessment (health care))[Mesh major topic] OR “quality indicators, health care”[Mesh major topic] OR quality indicator\*[ti] OR outcome indicator\*[ti] OR outcome measure\*[ti] OR adherence[ti] OR compliance[ti]) AND (((English[la] OR dutch[la]))) Filters: abstract available; publication date from 2006/12/01 to 2016/12/31; Humans

### Search for prognostic factors

(non-specific low back pain) AND (prognostic factors AND “last 5 years”[PDat]) (non-specific low back pain) AND (risk factors AND “last 5 years”[PDat])

Filter: review

## Appendix C

All measures identified process and outcome measures

Included outcome and process measures	Type of measure	Measure description	Derived from guideline or other scientific literature/indicator sets
Quebec Back Pain Disability Scale (QBPDS) [42]	Outcome	This is a 20-item questionnaire that measures functional limitations for low back pain	[27, 34, 43, 44]
Oswestry Disability Index (ODI) [45]	Outcome	The ODI is a 10-item questionnaire is designed to assess limitations of various activities of daily living	[9, 21, 34, 38, 46, 47]
Patient-Specific Functional Scale (PSFS) [48]	Outcome	Patients rate their ability to complete an activity on a 11-point scale at a level experienced prior to injury or change in functional status	[27, 43, 44, 49]
Numeric Pain Rating Scale (NPRS) [50]	Outcome	For pain intensity, the scale is most commonly anchored by ‘no pain’ (score of 0) and ‘pain as bad as it could be’ or ‘worst imaginable pain’ (score of 10)	[9, 27, 34, 38, 46, 47]
Global Perceived Effect—Dutch Version (GPE-DV) [51]	Outcome	This is a 2-item questionnaire about patient satisfaction and experienced effect of the treatment on a 7-point Likert scale	[27, 44]
Improvement of activities	Outcome	Derived from a quality indicator that scores the improvement in activities	[52]
Patient-Reported Experience Measures Physiotherapy (PREM-P) [53]	Outcome	This is a 17-item questionnaire for patient satisfaction about the physiotherapist, practice, and the received treatment	[53]
Subgroups for Targeted Treatment (STarT) Back Screening Tool (SBT) [39]	Screening tool	The SBT is a screening instrument for patients with non-specific low back pain based on patient characteristics and psychosocial factors	[21, 38, 47]
Tampa Scale of Kinesiophobia (TSK) [54]	Outcome	This questionnaire scores pain-related fear on a 17-item questionnaire for patients with low back pain	[44]
Number of treatment sessions	Outcome	Derived from a quality indicator that scores the improvement in activities	[43]
History taking described in the Electronic Health Record	Process	The percentage characteristics from the anamnesis (request for help, duration, and course of recovery) that is described in the EHR	[8, 55, 56]
Treatment plan described in the Electronic Health Record	Process	The percentage characteristics from the treatment plan that is described in the EHR	[8, 55, 56]
Instrument included	Type of measure	Content	Developed with guidelines
Prognostic profiles from guidelines	Prognostic profiles	These prognostic profiles are derived from several guidelines, but most were based on the Dutch physiotherapy guideline from the KNGF	Mainly [27] and compare [18, 21, 33, 34, 36–38]
<b>Excluded</b>			
SF-36, Short Form Health Survey; SF-12, Short Form Health Survey; EQ-5D-3L, EuroQol-5 Dimensions-3 level; EQ-5D, EuroQol-5 Dimensions; EQ-VAS, EuroQol-Visual Analogue Scale; FABQ, Fear-Avoidance Beliefs Questionnaire; PCS, Pain Catastrophizing Scale; HADS, Hospital Anxiety and Depression Scale; GHQ, General Health Questionnaire; BPI, Brief Pain Inventory; BDI, Beck Depression Inventory; STAI, State-Trait Anxiety Inventory; OMPQ, Örebro Musculoskeletal Pain Questionnaire; RMDQ, Roland-Morris disability questionnaire; VAS, Visual Analogue Scale			

## Appendix D

Summary of individual prognostic factors and an overview of useful components in the guidelines for a prognostic profile for NSLBP.

Individual prognostic factors	Prognostic profiles	Guideline
Depression, anxiety, catastrophizing	No classification based on prognostic profiles	American Physical Therapy Association [34]
Referred leg pain, comorbid pain, disability, catastrophizing, fear, anxiety, and depression	Described three profiles based on prognostic (psychosocial) factors: low, medium, and high risks	A primary care back pain screening tool [39]
Psychosocial factors are associated with the course of recovery	Derived from text: 0–12 weeks of acute NSLBP 12 < weeks of chronic NSLBP	CBO guideline [37]
Described psychosocial factors are important values in the course of recovery	No classification based on prognostic profiles	European guidelines for acute NSLBP [32]
No advice in work-related posture, the longer someone is absent on work, the longer the re-integration, psychosocial factors are delaying recovery, heavy physical work	No classification based on prognostic profiles	European guidelines for chronic NSLBP [33]
Many individual prognostic factors are described and focused on back pain-related factors, individual factors, work-related factors, psychosocial factors	Described three profiles based on prognostic (psychosocial) factors Normal course of recovery Abnormal course of recovery Abnormal course of recovery (with psychosocial factors)	KNGF guideline for low back pain [27]
Anxiety of movement, fear avoidance, catastrophizing, passive coping strategies	Described two profiles: Patient a < 12 weeks of recovery Patient b > 12 weeks of recovery	Ketenzorg guideline NSLBP [31]
Psychosocial factors are related to the course of recovery	Derived from text: 6 weeks of acute NSLBP 6–12 weeks of sub-acute NSLBP 12 < weeks of chronic NSLBP	NHG guideline [36]
Avoiding normal activities based on inappropriate beliefs about their condition	Advises the STarT Back Screening Tool (SBT)	NICE [38]
All guidelines described that psychosocial factors are related to the course of recovery	Derived from text most guidelines described: 0–6 weeks of acute NSLBP 6–12 weeks of sub-acute NSLBP 12 < weeks of chronic NSLBP	Overview of clinical guidelines for NSLBP [35]
Pain catastrophizing, fear of movement, anxiety, and depression	Three treatment profiles: Medical treatment Rehabilitation treatment Self-care management	Treatment-based classification (TBC) system [21]

## Appendix E

### Prognostic profile

Table 1 Prognostic profiles	Course	Prognostic risk factors	Prognostic psychosocial factors	Length of recovery
Profile A	Normal course of recovery	No risk factors (see Table 2)	Non-dominant psychosocial factors (see Table 2)	Less than 6 weeks
Profile B	Abnormal course of recovery	Risk factors often present (see Table 2)	Non-dominant psychosocial factors (see Table 2)	Less than 12 weeks
Profile C	Abnormal course of recovery	Risk factors often present (see Table 2)	Dominant psychosocial factors often present (see Table 2)	More than 12 weeks

Table 2 Examples of prognostic risk factors in NSLBP

Back pain-related factors	High level of restrictions during daily activities Spread pain down the leg
Individual factors	Elderly age Poor health status Episodes of NSLBP in the past
Work-related factors	No advice in work-related posture The longer someone is absent on work, the longer the re-integration Bad relationship with colleagues Heavy physical work
Prognostic psychosocial factors	Avoiding activities Catastrophizing ideas about pain Passive coping strategies Depression Somatization

In addition, on the guideline search we found 15 studies in our systematic search on PubMed and included 10 systematic reviews for selecting individual prognostic factors [57–67]. We found that catastrophizing and somatization seem to be negatively associated with recovery of NSLBP [58, 59]. Anxiety for moving is negatively associated with treatment benefits [59–62]. Patients with a high level of depression need more recovery time in comparison with people without depression [60, 62, 63]. Previous episodes of NSLBP may be associated with delayed recovery [61, 62, 64–67]. Most of these prognostic factors were already described in the KNGF guideline (somatization, anxiety for moving, depression) or in other guidelines.

## Appendix F

Themes, categories, and codes identified based on the interviews with patients

Theme	Categories	Codes
Patient satisfaction	Contribution to quality	Satisfaction combined with treatment effect is quality Satisfaction is a measure for quality Satisfaction is good for quality Satisfaction of a group patients could be a measure of quality
	No contribution to quality	Satisfaction of treatment effects is different for individual patients
	Treatment effect	Satisfaction is depending on the treatment effect Satisfaction is an aspect of the treatment effect Satisfaction is depending on the chosen treatment of the PT Patients can be satisfied even without a treatment effect
	Personal attention	Aspect of satisfaction is to feel comfortable during the treatment Satisfaction is depending on personal contact of PTs Satisfaction is to be treated respectful by PTs Satisfaction is that the PT is listening with care Satisfaction is that PT is willing to help Satisfaction is how PTs communicate with the patient Satisfaction is that the PTs communicate positively
Administration	Contribution to quality	Administration is helpful but needs to be transparent Administration can contribute but needs to be measured Report the treatment effect contributes Administration to show your accountability Administration is helpful With administration the PT can evaluate the treatment
	No contribution to quality	Besides administration the treatment process needs to be measured
	Communication	For a colleague during takeovers Helps a patient story during takeovers Useful for communication to external organizations
	Content	Needs to be short and brief Not at the expense of the treatment time Administration is a tool
Number of treatment sessions	Contribution to quality	Could be useful to detect outliers in a patient population Bad PTs will need more treatments than a good PT Could contribute to quality Is useful within a group with the same characteristics Could contribute to quality
	No contribution to quality	No relation with the number of pain Differs from individual patients

Theme	Categories	Codes
Transparency	Contribution to quality	Treatment outcomes of practices are useful to choose a caregiver Treatment outcomes are helpful to score the expertise of PT Feedback from other patients can help Report the results of outcome measurements
	No contribution to quality	Advice of doctor is more useful Be careful with patient reviews, to subjective
PROMs	Contribution to quality	Useful to score pain Helpful to evaluate the effect of the treatment Useful for diagnosis Only when the questionnaire is measured at the beginning and end of treatment Only at group level not for individual patients Helps to give a more accurate treatment Contributes to quality
	No contribution to quality	Pain differs between individual patients Too much influence of the individual patient (home exercises, work, listening to your body) Outcomes of PROMs is not representative
	Readability	Very clear Easy to read and understand In general easy to read You need to take time for PROMs
	Time	10–15 min max 15 min max A few minutes Around 10 min 5 min to short and 20 min to long
	Domain	Pain and functional problems in relation to the complaints Measured functional status is good Combination of physical and mental problems Psychological status needs to be monitored

PT physiotherapist, PROMs patient-reported outcome measures

## References

- Richardson W (2001) Crossing the quality chasm: a new health system for the 21st century. Institute of Medicine, pp 1–8
- Streiner DL, Norman GR, Cairney J (2015) Health measurement scales: a practical guide to their development and use. Oxford University Press, Oxford
- Meerhoff GA, van Dulmen SA, Maas MJM, Heijblom K, Nijhuis-van der Sanden MWG, Van der Wees PJ (2017) Development and evaluation of an implementation strategy for collecting data in a national registry and the use of patient-reported outcome measures in physical therapist practices: quality improvement study. *Phys Ther* 97(8):837–851
- Reuben DB, Tinetti ME (2012) Goal-oriented patient care: an alternative health outcomes paradigm. *N Engl J Med* 366(9):777–779
- van der Wees PJ, Zaslavsky AM, Ayanian JZ (2013) Improvements in health status after Massachusetts health care reform. *Milbank Q* 91(4):663–689
- Van Der Wees PJ, Nijhuis-Van Der Sanden MW, Ayanian JZ, Black N, Westert GP, Schneider EC (2014) Integrating the use of patient-reported outcomes for both clinical practice and performance measurement: views of experts from 3 countries. *Milbank Q* 92(4):754–775
- Verkerk EVM, van Dulmen S, van der Wees P, Terwee C, Beurskens S, de Boer D, Bakker C, Vajda I, Zuidgeest M (2017) PROM-toolbox. <https://www.zorginzicht.nl/kennisbank/Pagina/prom-toolbox.aspx>. Retrieved 1 Mar 2018
- Scholte M, Neeleman-van der Steen CW, van der Wees PJ, Nijhuis-van der Sanden MW, Braspenning J (2016) The reasons behind the (non)use of feedback reports for quality improvement in physical therapy: a mixed-method study. *PLoS ONE* 11(8):e0161056
- Clement RC, Welander A, Stowell C, Cha TD, Chen JL, Davies M, Fairbank JC, Foley KT, Gehrchen M, Hagg O, Jacobs WC, Kahler R, Khan SN, Lieberman IH, Morisson B, Ohnmeiss DD, Peul WC, Shonnard NH, Smuck MW, Solberg TK, Stromqvist BH, Hooff ML, Wasan AD, Willems PC, Yeo W, Fritzell P (2015) A proposed set of metrics for standardized outcome reporting in the management of low back pain. *Acta Orthop* 86(5):523–533
- Chiarotto A, Boers M, Deyo RA, Buchbinder R, Corbin TP, Costa LOP, Foster NE, Grotle M, Koes BW, Kovacs FM, Lin CC, Maher CG, Pearson AM, Peul WC, Schoene ML, Turk DC, van Tulder MW, Terwee CB, Ostelo RW (2018) Core outcome measurement

- instruments for clinical trials in nonspecific low back pain. *Pain* 159(3):481–495
11. Bombardier C (2000) Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. *Spine* 25(24):3100–3103
  12. Deyo RA, Battie M, Beurskens AJ, Bombardier C, Croft P, Koes B, Malmivaara A, Roland M, Von Korff M, Waddell G (1998) Outcome measures for low back pain research: a proposal for standardized use. *Spine* 23(18):2003–2013
  13. Deyo RA, Dworkin SF, Amtmann D, Andersson G, Borenstein D, Carragee E, Carrino J, Chou R, Cook K, DeLitto A, Goertz C, Khalsa P, Loeser J, Mackey S, Panagis J, Rainville J, Tosteson T, Turk D, Von Korff M, Weiner DK (2014) Focus article: report of the NIH task force on research standards for chronic low back pain. *Eur Spine J* 23(10):2028–2045
  14. Chiarotto A, Deyo RA, Terwee CB, Boers M, Buchbinder R, Corbin TP, Costa LOP, Foster NE, Grotle M, Koes BW, Kovacs FM, Lin C-WC, Maher CG, Pearson AM, Peul WC, Schoene ML, Turk DC, van Tulder MW, Ostelo RW (2015) Core outcome domains for clinical trials in non-specific low back pain. *Eur Spine J* 24(6):1127–1142
  15. Rutten GM, Degen S, Hendriks EJ, Braspenning JC, Harting J, Oostendorp RA (2010) Adherence to clinical practice guidelines for low back pain in physical therapy: do patients benefit? *Phys Ther* 90(8):1111–1122
  16. Hamilton AB, Yano EM (2017) The importance of symbolic and engaged participation in evidence-based quality improvement in a complex integrated healthcare system: response to “The science of stakeholder engagement in research”. *Trans Behav Med* 7(3):492–494
  17. Goodman MS, Sanders Thompson VL (2017) The science of stakeholder engagement in research: classification, implementation, and evaluation. *Trans Behav Med* 7(3):486–491
  18. Burton AK, Balagué F, Cardon G, Eriksen HR, Henrotin Y, Lahad A, Leclerc A, Müller G, van der Beek AJ, CBWGoGfPiLBP (2006) Chapter 2: European guidelines for prevention in low back pain: November 2004. *Eur Spine J* 15(Suppl 2):s136–s168
  19. Foster NE, Mullis R, Hill JC, Lewis M, Whitehurst DG, Doyle C, Konstantinou K, Main C, Somerville S, Sowden G, Wathall S, Young J, Hay EM (2014) Effect of stratified care for low back pain in family practice (IMPACT Back): a prospective population-based sequential comparison. *Ann Fam Med* 12(2):102–111
  20. Karran EL, McAuley JH, Traeger AC, Hillier SL, Grabherr L, Russek LN, Moseley GL (2017) Can screening instruments accurately determine poor outcome risk in adults with recent onset low back pain? A systematic review and meta-analysis. *BMC Med* 15(1):13
  21. Alrwaily M, Timko M, Schneider M, Stevans J, Bise C, Hariharan K, Delitto A (2016) Treatment-based classification system for low back pain: revision and update. *Phys Ther* 96(7):1057–1066
  22. Foster NE, Hill JC, O’Sullivan P, Hancock M (2013) Stratified models of care. *Best Pract Res Clin Rheumatol* 27(5):649–661
  23. Campbell SM, Braspenning J, Hutchinson A, Marshall MN (2003) Research methods used in developing and applying quality indicators in primary care. *BMJ* 326(7393):816–819
  24. Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lizaro P, van het Loo M, McDonnell J, Vader J, Kahan JP (2001) The RAND/UCLA appropriateness method user’s manual. RAND Corporation
  25. Guideline International Network (G-I-N) database. <https://www.g-i-n.net>. Retrieved 1 Dec 2016
  26. Physiotherapy Evidence Database (PEDro). <https://www.pedro.org.au/>. Retrieved 1 Dec 2016
  27. Staal J, Hendriks E, Heijmans M, Kiers H, Lutgers-Boomsma G, Rutten G (2013) KNGF-guideline for low back pain Verantwoording en toelichting. KNGF 21
  28. Jones J, Hunter D (1995) Consensus methods for medical and health services research. *BMJ Br Med J* 311(7001):376
  29. Cantrill J, Sibbald B, Buetow S (1996) The Delphi and nominal group techniques in health services research. *Int J Pharm Pract* 4(2):67–74
  30. Uscher-Pines L, Babin SM, Farrell CL, Hsieh YH, Moskal MD, Gaydos CA, Rothman RE (2010) Research priorities for syndromic surveillance systems response: consensus development using nominal group technique. *J Public Health Manag Pract JPHMP* 16(6):529–534
  31. Tulder MW, Custers JWH, de Bie R, Hammelburg R, Kolhaar B, Kuijpers T (2010) Nonspecific low back pain. CBO-guideline low back pain
  32. van Tulder M, Becker A, Bekkering T, Breen A, del Real MT, Hutchinson A, Koes B, Laerum E, Malmivaara A (2006) Chapter 3: European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J* 15:S169–S191
  33. Airaksinen O, Brox JJ, Cedraschi C, Hildebrandt J, Klüber-Moffett J, Kovacs F, Mannion AF, Reis S, Staal JB, Ursin H, Zanoli G (2006) European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J* 15:S192–S300
  34. Delitto A, George SZ, Dillen LV, Whitman JM, Sowa G, Shekelle P, Denninger TR, Godges JJ (2012) Low back pain. *J Orthop Sports Phys Ther* 42(4):A1–A57
  35. Koes BW, van Tulder M, Lin CW, Macedo LG, McAuley J, Maher C (2010) An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *Eur Spine J* 19:2075–2094
  36. NHG-standaard aspecifieke lagerugpijn (2005). <https://www.nhg.org/standaarden/volledig/nhg-standaard-aspecifieke-lagerugpijn>. Retrieved 1 Dec 2016
  37. CBO richtlijn aspecifieke lage rugpijn (2003). <http://www.kwaliteitskoepel.nl/assets/structured-files/NOV/Aspeciefekelagerugklachten2003.pdf>. Retrieved 1 Dec 2016
  38. NICE (2016) Low back pain and sciatica in over 16 s: assessment and management
  39. Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, Hay EM (2008) A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arth Rheum* 59(5):632–641
  40. Waddell G (2005) Subgroups within “nonspecific” low back pain. *J Rheumatol* 32(3):395–396
  41. Westby MD, Klemm A, Li LC, Jones CA (2016) Emerging role of quality indicators in physical therapist practice and health service delivery. *Phys Ther* 96(1):90–100
  42. Kopec JA, Esdaile JM, Abrahamowicz M, Abenhaim L, Wood-Dauphinee S, Lamping DL, Williams JI (1995) The Quebec back pain disability scale: measurement properties. *Spine* 20(3):341–352
  43. Calsbeek H, Maassen I, Braspenning JCC (2009) Publieke indicatoren eerstelijnszorg voor mensen met aspecifieke lage-rugpijn. Deel 8 uit de serie Voorhoedeproject Eerstelijns en Acute zorg – publieke indicatoren voor de eerstelijnszorg. IQ Healthcare, Nijmegen
  44. Van Dulmen SA, Schlieff A, Oostendorp RA, Nijhuis-Van Der Sanden MW (2009) Einrapport project kwaliteitsindicatoren Fysiotherapie-3 (ProKwaF-3) Kwaliteit fysiotherapeutische verslaglegging. IQ Healthcare
  45. Fairbank JC, Couper J, Davies JB, O’Brien JP (1980) The Oswestry low back pain disability questionnaire. *Physiotherapy* 66(8):271–273
  46. Dutch Spine Surgery Registry vragenlijsten patientfeedback versie 3.0 2017. Stichting DICA. <https://dica.nl/media/1255/DSSR%20Vragenlijsten%20V3.pdf>. Retrieved 1 Feb 2017
  47. van Hooff ML, van Loon J, van Limbeek J, de Kleuver M (2014) The Nijmegen decision tool for chronic low back pain.

- Development of a clinical decision tool for secondary or tertiary spine care specialists. *PLoS ONE* 9(8):e104226
48. Beurskens AJ, de Vet HC, Koke AJ, Lindeman E, van der Heijden GJ, Regtop W, Knipschild PG (1999) A patient-specific approach for measuring functional status in low back pain. *J Manip Physiol Ther* 22(3):144–148
  49. Bredero-Cohen AB et al (2009) VvOCM Richtlijn specifieke lage-rugklachten. <https://vvocm.nl/LinkClick.aspx?fileticket=4yIvPoam1WU%3D&tabid=178&portalid=1&mid=660>. Retrieved 1 Dec 2016
  50. McCaffery M (2001) Using the 0-to-10 pain rating scale. *Am J Nurs* 101(10):81–82
  51. Hudak PL, Wright JG (2000) The characteristics of patient satisfaction measures. *Spine* 25(24):3167–3177
  52. Driehuis F, Van Dulmen SA, Nijhuis-Van Der Sanden MW, Van Der Wees PJ (2016) Eindrapportage Nelson Beattie: een vragenlijst naar patiënt ervaringen in de fysiotherapie. IQ Healthcare
  53. Heeren J et al (2016) Ontwikkeling van de PREM Fysiotherapie; een vragenlijst over de ervaren kwaliteit van de fysiotherapeutische zorg. <https://www.nivel.nl/sites/default/files/bestanden/Ontwikkeling-PREM-Fysiotherapie.pdf>. Retrieved 1 Dec 2016
  54. Vlaeyen JW, Kole-Snijders AM, Boeren RG, van Eek H (1995) Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain* 62(3):363–372
  55. Heerkens Y, Hendriks H, de Graaf-Peters V (2011) KNGF guideline Fysiotherapeutische verslaglegging Nederlands. *Tijdschrift voor Fysiotherapie* 121:1–33
  56. van Dulmen SA, van Der Wees PJ, Meerhoff GA, Maas MJM, Driehuis F, Bakker-Jacobs A, Cruisberg J, Liefers J, Nijhuis-Van Der Sanden MW (2017) Eindrapport onderzoeksprogramma Masterplan Kwaliteit in Beweging (MKIB). IQ Healthcare, Nijmegen
  57. Fagundes FR, de Melo do Espirito Santo C, de Luna Teixeira FM, Tonini TV, Cabral CM (2017) Effectiveness of the addition of therapeutic alliance with minimal intervention in the treatment of patients with chronic, nonspecific low back pain and low risk of involvement of psychosocial factors: a study protocol for a randomized controlled trial (TalkBack trial). *Trials* 18(1):49
  58. Momsen AM, Jensen OK, Nielsen CV, Jensen C (2014) Multiple somatic symptoms in employees participating in a randomized controlled trial associated with sickness absence because of nonspecific low back pain. *Spine J Off J N Am Spine Soc* 14(12):2868–2876
  59. Wertli MM, Rasmussen-Barr E, Weiser S, Bachmann LM, Brunner F (2014) The role of fear avoidance beliefs as a prognostic factor for outcome in patients with nonspecific low back pain: a systematic review. *Spine J*
  60. George SZ, Beneciuk JM (2015) Psychological predictors of recovery from low back pain: a prospective study. *BMC Musculoskelet Disord* 16:49. <https://doi.org/10.1186/s12891-015-0509-2>
  61. Janwantanakul P, Sitthipornvorakul E, Paksaichol A (2012) Risk factors for the onset of nonspecific low back pain in office workers: a systematic review of prospective cohort studies. *J Manip Physiol Ther* 35(7):568–577
  62. Ramond-Roquin A, Bouton C, Begue C, Petit A, Roquelaure Y, Huez JF (2015) Psychosocial risk factors, interventions, and comorbidity in patients with non-specific low back pain in primary care: need for comprehensive and patient-centered care. *Front Med* 2:73
  63. Pinheiro MB, Ferreira ML, Refshauge K, Ordonana JR, Machado GC, Prado LR, Maher CG, Ferreira PH (2015) Symptoms of depression and risk of new episodes of low back pain: a systematic review and meta-analysis. *Arth Care Res* 67(11):1591–1603
  64. Harms MC, Peers CE, Chase D (2010) Low back pain: what determines functional outcome at six months? An observational study. *BMC Musculoskelet Disord* 11:236
  65. Helmhout PH, Staal JB, Heymans MW, Harts CC, Hendriks EJ, de Bie RA (2010) Prognostic factors for perceived recovery or functional improvement in non-specific low back pain: secondary analyses of three randomized clinical trials. *Eur Spine J* 19:650–659
  66. Hill JC, Fritz JM (2011) Psychosocial influences on low back pain, disability, and response to treatment. *Phys Ther* 91(5):712–721
  67. Ramond A, Bouton C, Richard I, Roquelaure Y, Baufreton C, Legrand E, Huez JF (2011) Psychosocial risk factors for chronic low back pain in primary care: a systematic review. *Fam Pract* 28(1):12–21. <https://doi.org/10.1093/fampra/cmq072>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

A. C. Verburg<sup>1</sup>  · S. A. van Dulmen<sup>1</sup> · H. Kiers<sup>2,3</sup> · M. W. G. Nijhuis-van der Sanden<sup>1</sup> · P. J. van der Wees<sup>1</sup>

✉ A. C. Verburg  
koen.verburg@radboudumc.nl

<sup>1</sup> Radboud University Medical Center, Radboud Institute for Health Sciences, IQ Healthcare, Nijmegen, The Netherlands

<sup>2</sup> Institute for Human Movement Studies, Utrecht University of Applied Sciences, Utrecht, The Netherlands

<sup>3</sup> Association for Quality in Physical Therapy (SKF), Zwolle, The Netherlands