

Towards an ICF- and IMMPACT-Based Pain Vocational Rehabilitation Core Set in the Netherlands

M. F. Reneman · T. T. Beemster · M. J. A. Edelaar ·
J. M. van Velzen · C. van Bennekom ·
R. Escorpizo

Published online: 2 February 2013
© Springer Science+Business Media New York 2013

Abstract *Background* For clinical use and research of pain within the context of vocational rehabilitation, a specific core set of measurements is needed. The recommendations of the International Classification of Functioning, Disability and Health (ICF) brief Core Set for Vocational Rehabilitation (VR) and those of Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) cover two broad areas. These two sources can be integrated when made applicable to vocational rehabilitation and pain. *Objective* To develop a core set of diagnostic and evaluative measures specifically for vocational rehabilitation of patients with subacute and chronic musculoskeletal pain, while using the brief ICF core set for VR as the reference framework in VR, and the IMMPACT recommendations in the outcome measurements around pain. *Methods* Three main steps were taken. The first step was to remove irrelevant and duplicate domains of the brief ICF Core Set for Vocational

Rehabilitation and the IMMPACT recommendations around pain. The second step was to match the remaining domains with existing instruments or measures. Instruments were proposed based on availability and its proven use in Dutch practice and based on proof of sufficient clinimetric properties. In step 3, the preliminary VR-Pain core set was presented to 3 expert panels: proposed users, Dutch pain rehabilitation experts, and international VR experts. *Results* Experts agreed with the majority of the proposed domains and instruments. The final VR-Pain Core Set consists of 18 domains measured with 12 instruments. All instruments possessed basic clinimetric properties. *Conclusion* An agreed-upon VR-Pain Core Set with content that covers relevant domains for pain and VR and validated instruments measuring these domains has been developed. The VR-Pain Core Set may be used for regular clinical purposes and research in the field of

M. F. Reneman (✉) · T. T. Beemster
Department of Rehabilitation Medicine, Center
for Rehabilitation, University Medical Center Groningen,
University of Groningen, 9750 RA Haren, P.O. Box 30.002,
Groningen, The Netherlands
e-mail: m.f.reneman@cvr.umcg.nl

T. T. Beemster · M. J. A. Edelaar · J. M. van Velzen ·
C. van Bennekom
Department of Research and Development, and Institute
of Vocational Assessment and Education, Heliomare
Rehabilitation Center, Wijk aan Zee, The Netherlands

T. T. Beemster · J. M. van Velzen
Coronel Institute of Occupational Health, Academic Medical
Center, University of Amsterdam, Amsterdam, The Netherlands

R. Escorpizo
Swiss Paraplegic Research (SPF), Nottwil, Switzerland

R. Escorpizo
ICF Research Branch in cooperation with the WHO
Collaborating Centre for the Family of International
Classifications in Germany (at DIMDI), Nottwil, Switzerland

R. Escorpizo
Department of Health Sciences and Health Policy, University
of Lucerne and SPF, Nottwil, Switzerland

R. Escorpizo
Department of Physical Therapy, Louisiana State University
Health Sciences Center, New Orleans, LA, USA

vocational rehabilitation and pain, but adaptations should be considered for use outside the Netherlands.

Keywords ICF · IMMPACT · Musculoskeletal pain · Vocational rehabilitation · Work rehabilitation · Employment · Return to work

Introduction

Chronic musculoskeletal pain has a substantial negative impact on quality of life and the ability to engage in meaningful activities and participation in the society, including work [1, 2]. In the Netherlands, musculoskeletal disorders such as back, neck and shoulder pain constitute about 35 % of all sickness absence and long-term disability compensations [3–5]. Medical care utilization and sickness absence due to musculoskeletal pain are associated with high economic burden to society similar to other western countries worldwide [6]. The majority (~80 %) of the costs are related to the inability to work [5]. One of the preferred interventions to promote return to work for patients with chronic musculoskeletal pain is vocational rehabilitation, because it has been proven to be effective in reducing disability and improving work participation, and it appears to be cost-effective [7, 8]. To further improve the effectiveness of vocational rehabilitation to optimize work participation, it is recommended to intervene as soon as possible, perhaps even as soon as the sub-acute phase of musculoskeletal pain [9, 10]. A network of 14 rehabilitation centers in the Netherlands has been established to deliver evidence-based vocational rehabilitation for workers with sub-acute and chronic musculoskeletal pain.

Vocational rehabilitation (VR) in its broadest form has recently been defined in a position paper as ‘a multidisciplinary evidence-based approach that is provided along a continuum of services and activities to working age individuals with health-related impairments, limitations, or restrictions with work functioning, and whose primary aim is to optimize work participation [11]. The authors of the position paper proposed the use of the International Classification of Functioning, Disability and Health (ICF) within the VR field (regardless of health condition). On one hand, the ICF Core Set for Vocational Rehabilitation has been developed with two versions: (1) the 90 ICF categories of the comprehensive version is intended for multidisciplinary setting and (2) the 13 ICF categories of the brief version is intended for single discipline encounter or clinical trials. The brief version due to less number of ICF categories is doable for practical application and feasible in VR-related patient evaluation and assessment [12]. On the other hand, the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT), has provided

recommendations for improving the design, execution, and interpretation of clinical trials of treatments specific to patients with pain [13, 14]. IMMPACT has proposed a core set of outcome measures for patients with pain. These recommendations were also broad, because they were intended to encompass the wide field of pain care, which extends far beyond the context of vocational rehabilitation.

So, here are two sets of recommendations or sets of domains, ICF Core Set and IMMPACT, which broadly address VR and pain, respectively. For clinical use and research of pain within the context of VR, developing a specific core set is needed, while learning from the two existing sets. The recommendations of the brief ICF Core Set for Vocational Rehabilitation and those of IMMPACT cover two broad areas, but should be merged and made applicable to a specific context, patient group and setting (pain and VR in the Netherlands in this study). However, we did not find papers relevant to the Netherlands, or anywhere else, describing the process and outcome of an ICF-IMMPACT core set, let alone the operationalization of those domains. The aim of the present study was to develop a core set of diagnostic and evaluative (clinical and economic) measures specifically for vocational rehabilitation of patients with sub-acute and chronic musculoskeletal pain, while using the ICF as the reference framework in VR, and IMMPACT in the outcome measurements around pain. In this study, the context is situated in the Netherlands, including its health care and social security policies as of the year 2012. As part of integrating our knowledge on the ICF, work, and pain, our research question is: how can the brief version of the ICF Core Set for Vocational Rehabilitation and the IMMPACT recommendations be best applied in one blended VR-Pain Core Set for patients with sub-acute and chronic musculoskeletal pain in the Netherlands?

Methods

Three main steps were taken. The first step was to remove irrelevant and duplicate domains of the brief ICF Core Set for Vocational Rehabilitation and the IMMPACT recommendations. Irrelevant domains were defined as those domains that do not apply or only apply to an estimated 1 % of the target population (as judged by the authors and the expert panel). Duplicate domains were defined as domains that cover overlapping, equal or very similar content or concept. Additionally, the remaining domains were checked to see whether they could be used for economic evaluations also. If not, this was added. The second step was to match the remaining domains with existing instruments or measures. Instruments were proposed based on availability and its proven use in Dutch practice and peer reviewed literature. Existing instruments were

included based on proof of sufficient reliability (test–retest reliability: Intra Class Coefficient (ICC) >0.90 (preferred), Kappa >0.60, Pearson correlation coefficient >0.80; internal consistency: Cronbach’s alpha >0.80 [15]; construct validity (yes/no/not applicable (na)); responsiveness to change (yes/no/na; relevant for outcome measures only); existence of a validated version in Dutch language (yes/no; relevant for questionnaires only); and feasibility (acceptable patient and practitioner burden: yes/no). The second step was not performed to provide a systematic review of the psychometric properties of all instruments available, but to check whether the psychometric properties of the proposed instruments of the preliminary VR-Pain Core Set were acceptable.

The result of step 1 and 2 was a preliminary version of what we would call the VR-Pain core set. In step 3, to be informed by input from relevant people, the preliminary VR-Pain core set was presented to 3 expert or user panels: Dutch VR centers (proposed users (management and clinicians); n = 13), Dutch pain rehabilitation development centers (pain rehabilitation experts; n = 4), and members of the VR-Pain Core Set consensus group (VR experts; n = 23) [12]. Participants were sent the introduction to, methods and results of steps 1 and 2, including the preliminary VR-Pain core set. They were asked whether they agreed with the taken steps and the proposed core domains of the preliminary VR-Pain Core Set, and whether they agreed with the proposed instruments. In case of non-agreement, they were asked to explain their disagreement and to suggest improvements. In case the comments were unclear, the first author contacted the responder. All participants had 3 weeks to respond. Participants were sent a reminder after 2 weeks. The authors of this paper then synthesized the comments of the responders into a final VR-Pain core set.

Results

Step 1

The domains of the brief VR-Pain Core Set and the IMMPACT recommendations are presented in Tables 1 and 2. The results of Step 1, the selection of irrelevant domains and reduction of duplicates, are also presented in Tables 1 and 2.

Step 2

Results of the process of matching core set domains to instruments, including its quality appraisal, are presented in Table 3. Additions as described in step 3 were also incorporated in Table 3. Domains from the IMMPACT

recommendations were provided with ICF codes, with exception of personal factors which are not currently coded in the ICF.

Step 3

The preliminary VR-Pain core set was emailed to members of the expert panels in February 2012. Overall response was n = 18 (response rate 45 %); proposed users n = 11 (85 %), pain rehabilitation experts n = 4 (100 %), VR experts n = 3 (13 %). Of the VR experts, an additional n = 3 responded that the specific nature of the subject of this study was out of their field of expertise. One of the VR experts was contacted by phone, because the answers and comments were ambiguous. Eleven (61 %) respondents agreed with the proposed domains of the preliminary core set, while five disagreed, and two did not answer. Ten (55 %) respondents agreed with the proposed instruments of the preliminary core set, five disagreed, and three did not answer or indicated to have insufficient knowledge to judge. ‘Disagreements’ were most often accompanied by a short explanation and/or suggestion. The project members have decided unanimously that some comments should not be regarded as disagreements with the proposed domains or instruments, but rather as an item that a single expert proposed to add to the preliminary set. However, because not single experts, but rather the brief ICF Core Set for Vocational Rehabilitation and the IMMPACT recommendations formed the basis of this new and specific core set, it was decided that items proposed by single experts were not added to the definitive set, unless the project team decided otherwise based on the underlying core sets.

Based on the responses of the participants, the following domains were added to the VR-Pain Core set: adverse effects that has not lead to discontinuation of the program (adherence to the intervention; treatment records) and personal problems unrelated to work (Work Reintegration Questionnaire; WRQ). With regard to the instruments, the following measurements were changed or added: energy and drive functions (ICF code b130) will be measured with numerical rating scale (NRS) for fatigue; physical functioning will not be measured with the Pain Disability Index only, but also with RAND-36 scale physical functioning; Astrand or Bruce submaximal ergometry will be used to measure exercise tolerance functions; assessment of functioning at home or in unpaid work will be added as part of the demographic questionnaire.

Description of Instruments of the Final VR-Pain Core Set

The *EuroQol-5D* (EQ-5D) is a 6-item questionnaire to investigate quality of life. The EQ-5D categories measure 5 dimensions: mobility, self-care, activities of daily life, pain

Table 1 ICF categories of the brief ICF Core Set for Vocational Rehabilitation and relevance of the domains to the proposed VR-Pain core set

ICF code	ICF category title	Relevant	Comments
<i>Activities and participation</i>			
d155	Acquiring skills	No	This is not a key challenge in patients with pain. This item was included in the brief VR-Pain Core Set to accommodate individuals with neurological diagnoses and intellectual and cognitive challenges
d240	Handling stress and other psychological demands	Yes	
d720	Complex interpersonal interactions	Yes	
d845	Acquiring, keeping and terminating a job	No	The target population is employed. Aim of VR in our case is to return to own work and same employer, or to improve work performance. Keeping a job: duplicate concept with d850
d850	Remunerative employment	Yes	Work status will be assessed, including absenteeism and presenteeism
d855	Non-remunerative employment	No	Only patients with paid work are admitted to our specified setting
<i>Environmental factors</i>			
e310	Immediate family	Yes	
e330	People in positions of authority	Yes	
e580	Health services, systems and policies	No	Within the target population, this item is of relevance, but not variable across subjects in the Netherlands
e590	Labour and employment services, systems and policies	No	Within the target population, this item is of relevance, but not variable across subjects in the Netherlands
<i>Body functions</i>			
b130	Energy and drive functions	Yes	
b164	Higher-level cognitive functions	No	Within this target population, high-level cognitive functions are unaffected
b455	Exercise tolerance functions	Yes	

and anxiety/depression. Five questions are categorical (1–3 scale) and one question assessing general health status is on interval level (VAS 0–100). A Dutch language version of the EQ-5D is available [16, 17]. The EQ-5D is a widely employed instrument to assess health related quality of life (QoL), is used in cost effectiveness research based on quality adjusted life years (QALY) and is recommended by the Dutch Healthcare Insurance Board [18]. Lower levels of QoL are associated with productivity loss in patients with low back pain [19].

A single item of the *Work Ability Index* (WAI) will be used to assess self-reported work ability. Current work ability compared to lifetime best can be scored on a 0–10 response scale, where 0 represents ‘completely unable to work’ and 10 ‘work ability at its best’. A strong association between the single item and the complete WAI was observed ($r = 0.87$) [20].

The *PROductivity and DISease Questionnaire* (PRO-DISQ) [21] will be used to assess employment status, absenteeism and presenteeism. Absenteeism refers to time off from work. Presenteeism refers to productivity loss while at-work. Both may be associated with a health condition. Absenteeism is measured with a three-month recall period, and will be measured specifically related to pain condition. The number and duration of a maximum number

of three absenteeism periods are collected. Presenteeism is measured with two items on a 11-point scale, also known as the QQ-index (quantity and quality). The first item measures quality of work done in the last day at work, ranging from 0 (I couldn’t do anything) to 10 (I could do the same as normal). The second item measures quantity of work done in the last day at work, ranging from 0 (the quality of my work was dramatic) to 10 (the quality of my work was normal).

The *Pain Disability Index* (PDI) is a 7-item questionnaire to investigate the magnitude of the self-reported pain related disability, independent from region of pain or pain-related diagnosis. The questionnaire is constructed on a 0–10 numeric rating scale in which 0 means no disability and 10 maximum disability. Total scores can range from 0 to 70, with higher scores reflecting higher interference of pain with daily activities. The PDI measures family / home responsibilities, recreation, social activity, occupation, sexual behaviour, self-care and life support activity [22, 23].

The *RAND-36 scale physical functioning* will be used to measure self-reported physical functioning independent of (pain) diagnosis [24]. The RAND-36 has been used widely across health conditions (www.rand.org, accessed August 2012). The physical functioning scale consists of 10

Table 2 IMMPACT recommendations and supplemental domains and relevance of the domains to the proposed VR-Pain core set

	Relevant	Comments
<i>Core domains</i>		
Pain	Yes	
Physical functioning	Yes	
Emotional functioning	Yes	
Participant ratings of global improvement	Yes	
Symptoms and adverse events	Yes	Symptoms duplicate with pain. Adverse events will be monitored under participant disposition
Participant disposition (including adherence to the treatment regimen and reasons for premature withdrawal from the trial)	Yes	Will be replaced by: Adherence to the intervention and reasons for premature withdrawal
<i>Supplemental domains</i>		
Role functioning (i.e. work and educational activities)	Yes	
Interpersonal functioning (i.e. relationships and activities with family, friends, and others)	Yes	Duplicate. Will be covered under immediate family and people in authority (as mentioned in the ICF-VR), which are the primary group of interest in our context
Pharmacoeconomic measures and health care utilization	Yes	Will be included as one domain: health care utilization
Biological markers (e.g. assessments based on quantitative sensory testing, imaging, genetic markers, pharmacogenomics, and punch skin biopsy)	No	The target population includes patients with non-specific pain. If biological functions are relevantly involved in the health status, patients are excluded because this could indicate a specific pain syndrome
Coping	Yes	
Clinician or surrogate ratings of global improvement	Yes	
Neuropsychological assessments of cognitive and motor function	Yes	Duplicate. Will be covered under coping / stress and psychological demands and exercise tolerance and physical functioning, all part of ICF-VR
Suffering and other end-of-life issues	No	Not applicable for the target population

questions with 3 possible answers on a Likert scale: ‘yes, strongly limited’, ‘yes, a bit limited’, and ‘no, not limited’. The total score can range from 0 to 100, with higher scores indicating better physical functioning. The validity and reliability of the Dutch version are good [25].

The *Work Reintegration Questionnaire* (WRQ) is an instrument for assessing the most important psychosocial factors in the delay of recovery and work resumption. The questionnaire consists of 78 items distributed among 8 scales: ‘Distress’, ‘Illness behaviour’, ‘Job strain’, ‘Job dissatisfaction’, ‘Control’, ‘Avoidance’, ‘Perfectionism’ and ‘Stressful home situation’. The *Work Reintegration Questionnaire* (WRQ) measures the following dimensions: distress, interference, work stress, work satisfaction, insecurity / avoidance, perfectionism / persistence, home situation [26]. The questionnaire was developed in Dutch (VAR: vragenlijst arbeidsreintegratie). A validated translation in English is currently in development (personal communication with author).

Pain intensity and fatigue can be assessed using an 11-point NRS (*NRS-pain* and *NRS-fatigue*), ranging from 0 (no pain / fatigue) to 10 (worst possible pain / fatigue), requiring patients to rate their current and average intensity of the last 7 days [9].

Exercise tolerance functions will be assessed with standardized *lifting capacity tests* from the Workwell

functional capacity evaluation (FCE): lifting low and overhead lifting. Procedures are described in detail elsewhere [27]. These tests were found to be predictive of functional capacity performance in general in patients with back pain and neck / upper extremity pain [28]. A standardized submaximal *Astrand bicycle test* [29, 30] or *Bruce treadmill test* [31] will be used to assess exercise tolerance functions as well as energy and drive functions.

The *Trimbos iMTA questionnaire for measuring Costs of Psychiatric Illnesses* (TiC-P), module 1, will be used to assess health care utilization. The questionnaire has a recall period of 4 weeks. Visits and consultations of the following health care providers were measured: general practitioner, physiotherapist, manual therapist, exercise therapist, occupational therapist, psychologist, insurance physician, medical specialists in hospitals, hospitalization (number of days), occupational physician, social worker, and dietician. Further items were alternative care, home care, medication use, and job related care like job coaches, ergonomic changes at the work site and re-integration specialists [32]. Slight adaptations in the context and scope of health care practitioners were made to better fit TiC-P to the target population (i.e. from psychiatry to pain and work).

Table 3 Quality appraisal of VR-Pain core domains classified per ICF category

VR-Pain core domains	ICF code	Name of instrument or scale	Reliability	Construct validity	Responsiveness	Formally validated translation	Utility (min)	References
<i>Quality of Life^a</i>								
Quality of life	x	EQ-5D	Yes	Yes	Yes	Yes	2	[33]
<i>Activities/participation</i>								
Remunerative employment	d850	WAI q3	Yes	Yes	AoE	Yes	<1	[20]
		PRODISQ	Yes	Yes	AoE	Yes	5	[21]
Role functioning (i.e. work and educational activities)	d850	PDI q4	Yes	Yes	Yes	Yes	<1	[23, 34]
Physical functioning	d899	PDI total	Yes	Yes	Yes	Yes	2	[23, 34][24]
		RAND-36 physical functioning						[25]
Complex interpersonal interactions	d720	WRQ satisfaction	Yes	Yes	NA	Yes	1	[26]
Handling stress and other psychological demands	d240	WRQ work stress	Yes	Yes	NA	Yes	1	[26]
<i>Body functions</i>								
Pain	b280	Diagnosis	–	–	–	–	–	–
Pain intensity	b280	NRS pain	Yes	Yes	Yes	Yes	<1	[35, 36]
Energy and drive functions	b130	NRS fatigue	Yes	Yes	Yes	No	<1	[40]
Exercise tolerance functions	b455	Lifting test	Yes	Yes	AoE	NA	10	[28, 37][27, 38]
		Astrand bicycle ergometry	Yes	Yes	AoE	NA	15	[29, 30]
		Bruce treadmill ergometry	Yes	Yes	AoE	NA	10	[39]
Emotional functioning	b152	WRQ distress	Yes	Yes	NA	Yes	1	[26]
<i>Environmental factors</i>								
Immediate family	e310	PDI q1	Yes	Yes	NA	Yes	<1	[23, 34]
		WRQ home	Yes	Yes	NA	Yes	1	[26]
People in positions of authority	e330	WRQ satisfaction	Yes	Yes	NA	Yes	1	[26]
Health care utilization	e580	TiC-P	Yes	Yes	NA	Yes	8	[32]
<i>Personal factors</i>								
Coping	IM	WRQ avoidance	Yes	Yes	NA	Yes	1	[26]
		WRQ persistence	Yes	Yes	NA	Yes	1	[26]
<i>Evaluation</i>								
Participant ratings of global improvement	IM	GPE	Yes	Yes	Yes	Yes	<1	[23]
Adherence to the intervention and reasons for premature withdrawal; diagnosis;	IM	Medical records	AoE	AoE	AoE	NA	<1	–
Clinician or surrogate ratings of global improvement	IM	GPE	AoE	AoE	AoE	NA	<1	–

^a Not in IMMPACT or VR-Pain Core Sets; x ICF code not available; IM IMMPACT; EQ-5D EuroQol-5D; WAI work ability index; PRODISQ PROductivity and DISease Questionnaire; PDI pain disability index; WRQ work reintegration questionnaire; GPE global perceived effect; NA not applicable; AoE absence of evidence; TiC-P Trimbos iMTA questionnaire for measuring costs of psychiatric illnesses

Global perceived effect (GPE) can be measured with a 7-point Likert scale ranging from 1 to 7 (1; ‘extremely worsened’, 2; ‘much worsened’, 3; ‘little worsened’, 4; ‘unchanged’, 5; ‘little improved’, 6; ‘much improved’, 7; ‘completely improved’). Two GPE questions are proposed: how much did your treatment change your pain compared

to pre-treatment level, and how much did your treatment change your work status compared to pre-treatment level?

Treatment records will be used to assess diagnosis, adherence to the treatment program, adverse effects that has not lead to discontinuation of the program, and reasons for premature withdrawal.

Discussion

Sub-acute or chronic musculoskeletal pain can lead to a decrease in work participation up to the point where VR becomes essential. It is important to understand and address musculoskeletal pain in the context of VR because if we can mitigate the burden of work disability, we can facilitate early and sustained return to work. To do so, in this study, we attempted to blend two sources of domains around pain and VR, listed the instruments by which we can operationalize the domains, and developed a VR-Pain core set that may benefit clinical and research application in the VR-pain field in the Netherlands and potentially in other countries as well. To this end, the final VR-Pain core set consisted of 12 instruments that covered 18 domains.

As presented in Tables 1 and 2, domains and items of the underlying sets were removed by the authors because they were not deemed relevant for the (vast majority of) the target population. While this was not or incidentally challenged by members of the expert groups, this still needs elucidation. Acquiring skills (ICF code d155 *Acquiring skills*) for example, was excluded, because this it is not primarily affected (or core) in people with pain, and therefore not a goal in VR. *Acquiring skills* was deemed very relevant for the brief ICF Core Set for VR, because VR in its broadest form is provided to workers with a range of disabilities reaching far beyond pain, including workers with neurological and intellectual problems. Because of the specific setting for which the VR-Pain core set was developed, other items that were initially included to accommodate the wide application of both ICF and IMMPACT sets did not make to the core and final VR-pain set like unpaid work, acquiring, keeping, and terminating work, end-of-life issues, and higher level cognitive functions. To exclude the latter, however, may be subject to debate, because workers with pain often report challenges with concentration and memory. These concomitant complaints are regarded as related to pain and fatigue (which are already included in the final set), and perhaps symptoms related to central sensitization. Non-specific pain does not directly affect the brain and higher neurological functions as captured in ICF code b164 *Higher level cognitive functions*. For similar reasons biological markers were also excluded. Finally, while environmental issues such as insurance and social security systems are considered

relevant [4] and vary across jurisdictions, they do not vary across the workers in the specific setting for which this VR-Pain core set was developed for. For generalizations beyond the Dutch borders, we advise researchers to describe the issues in future reports within the context or controlling for insurance and social security systems.

Most instruments proposed to make up the VR-Pain core set all comply with basic clinimetric properties as presented in Table 3. Because the properties included validated Dutch language versions of questionnaires and feasibility, the set of instruments proposed is likely to differ from core sets for different countries and languages. Additionally, we have attempted to choose instruments that could be used for clinical as well as for research purposes, including economic evaluations. Even though EQ-5D was not recommended by either one of the underlying core sets (ICF Core Set or IMMPACT), it was added because this instrument can be used for economic evaluations. Additionally, it captures an important secondary aim of vocational rehabilitation, which is to contribute to increase quality of life.

New core sets that apply to specific groups may thus be developed based on existing core sets. The exercise of developing a new core set based on two established ones has not been presented previously. Thus, the methodology described in this paper is new. We have aimed to describe this methodology transparently, to enable readers to either replicate these steps when developing or validating other core sets specific to their setting, or to use it as a basis for further development of this methodology. By asking Dutch experts in the pain rehabilitation field, prospective users and international VR experts, we aimed to test the content validity of the newly developed core set. However, this paper may also be regarded as external validation of the underlying core sets. In choosing the instruments, we aimed to combine sound psychometric properties with the options for future cost-effectiveness studies or intervention trials. This will enable future users to study clinical and economic outcomes in the (Dutch) usual care setting, which should make a significant contribution to the field of VR and pain.

While the response rates of the Dutch pain rehabilitation experts and prospective users was high, response of the international experts was low. Some international experts responded that this exercise was specifically not in their sub-field of expertise (e.g. cognitive vocational rehabilitation), the majority of this group did not respond at all, which may be attributed to lack of time availability or were unable to follow up on the electronic invitation and reminder. The relevance of this non-response is unknown. Because based on the responses only small changes were made to the final core set, and no differences in response patterns between expert groups were observed, we assume that the relevance of the non-response to be limited. Patients were not invited to participate in this specific

exercise, because patient involvement was already incorporated in the development of the two underlying core sets. Both the ICF Core Set for Vocational Rehabilitation and the IMMPACT recommendations are in principle experts-based. Even though the VR-Pain core set was agreed upon by most experts ('externally validated'), it is in its essence also an experts-based core set. Future use will discover whether the set of instruments is deemed too extensive for routine clinical use, and where and why this core set should be adapted to new developments in the VR-Pain field.

In line with the recommendations underlying the ICF VR expert group [12], the lack of classification of the personal factors in the ICF which can play a crucial influence on work functioning, will need careful consideration in the future. Although some performance-based instruments are included in the final VR-Pain core set, the majority of the instruments are self-report based. Apart from its strengths (outcomes are judged by the patients, not by or interpreted by others), this may also introduce a risk of bias, particularly in the estimation of absenteeism and presenteeism. Additionally, while clinimetric properties of the individual instruments in the VR-Pain core set were checked, they were not systematically reviewed.

The clinical relevance of using this VR-Pain core set is that it will provide a firm base for routine clinical use and evaluation of services in vocational rehabilitation settings with pain-related cases. Clinicians can, based on their clinical expertise or professional guidelines, add diagnostic instruments to this core set as needed. Moreover, the VR-Pain core set should not replace clinical expertise, but rather should complement it. The methodology described in this paper may be generalizable to develop other setting-specific core sets or a combination thereof. Additionally, most of the instruments in the VR-Pain core set are used internationally, which will address generalizability and comparability. Costs calculations underlying the EQ-5D, PRODISQ and TiC-P questionnaires, however, are based on Dutch guidelines which are expected to be different from other countries. While the VR-Pain core set is developed for the Netherlands, the burden of pain and work disability in the Netherlands is similarly high as in other industrialized countries [1]. Therefore, it is recommended that similar core sets are to be developed and tested for different countries. To enable generalization across countries, facilitate common language and stimulate future developments, we recommend that whenever possible, the same instruments are used.

Conclusion

A VR-Pain core set with content that covers relevant domains for pain and VR and with validated corresponding instruments that measure these domains has been

developed. The VR-Pain core set may be used for clinical purposes, and (cost) effectiveness research in the field of vocational rehabilitation and pain. Caution is warranted for direct use outside of The Netherlands, because differences in cultural and service and political systems exist, hence the basis for costs calculations may be different. Additionally, for use and generalization beyond that of The Netherlands, it is recommended that environmental factors (ICF e580 and e590, Table 1) be considered and examined.

Acknowledgments The authors wish to acknowledge the input of all responders. The following responders have granted permission to acknowledge them: Pain rehabilitation experts: Karlein Schreurs, Rita Schiphorst Preuper, Albere Köke, Sylvia Remerie. Prospective users: Peter van der Wurff, Martin Smeulers, Loes Swaan, Levijn Romp, Claire Tilmans, Sieger de Vries, Marleen ter Haar. VR experts: Alex Burdorf, Debra Homa, Sven-Uno Marnetoft.

References

- Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. *Eur J Pain*. 2006;10(4):287–333.
- Reid KJ, Harker J, Bala MM, Truyers C, Kellen E, Bekkering GE, et al. Epidemiology of chronic non-cancer pain in Europe: narrative review of prevalence, pain treatments and pain impact. *Curr Med Res Opin*. 2011;27(2):449–62.
- Bekkering GE, Bala MM, Reid K, Kellen E, Harker J, Riemsma R, et al. Epidemiology of chronic pain and its treatment in The Netherlands. *Neth J Med*. 2011;69(3):141–53.
- Anema JR, Schellart AJ, Cassidy JD, Loisel P, Veerman TJ, van der Beek AJ. Can cross country differences in return-to-work after chronic occupational back pain be explained? An exploratory analysis on disability policies in a six country cohort study. *J Occup Rehabil*. 2009;19(4):419–26.
- Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine J*. 2008;8(1):8–20.
- Hoy D, March L, Brooks P, Woolf A, Blyth F, Vos T, et al. Measuring the global burden of low back pain. *Best Pract Res Clin Rheumatol*. 2010;24(2):155–65.
- Norlund A, Ropponen A, Alexanderson K. Multidisciplinary interventions: review of studies of return to work after rehabilitation for low back pain. *J Rehabil Med*. 2009;41(3):115–21.
- Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, et al. Chapter 4 European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J*. 2006;15(Suppl 2):S192–300.
- Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. *Pain*. 1986;27(1):117–26.
- Jensen IB, Busch H, Bodin L, Hagberg J, Nygren A, Bergstrom G. Cost effectiveness of two rehabilitation programmes for neck and back pain patients: a seven year follow-up. *Pain*. 2009;142(3):202–8.
- Escorpizo R, Reneman MF, Ekholm J, Fritz J, Krupa T, Marnetoft SU, et al. A conceptual definition of vocational rehabilitation based on the ICF: building a shared global model. *J Occup Rehabil*. 2011;21(2):126–33.
- Finger ME, Escorpizo R, Gassel A, Gmunder HP, Luckenkemper M, Chan C, et al. ICF core set for vocational rehabilitation: results of an international consensus conference. *Disabil Rehabil*. 2012;34(5):429–38.

13. Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2003;106(3):337–45.
14. Turk DC, Dworkin RH, Revicki D, Harding G, Burke LB, Cella D, et al. Identifying important outcome domains for chronic pain clinical trials: an IMMPACT survey of people with pain. *Pain*. 2008;137(2):276–85.
15. Wind H, Gouttebauge V, Kuijer PP, Frings-Dresen MH. Assessment of functional capacity of the musculoskeletal system in the context of work, daily living, and sport: a systematic review. *J Occup Rehabil*. 2005;15(2):253–72.
16. Lamers LM, McDonnell J, Stalmeier PF, Krabbe PF, Busschbach JJ. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. *Health Econ*. 2006;15(10):1121–32.
17. Lamers LM, Stalmeier PF, McDonnell J, Krabbe PF, van Busschbach JJ. Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff. *Ned Tijdschr Geneesk*. 2005;149(28):1574–8.
18. Bieleman HJ, Reneman MF, van Ittersum MW, van der Schans CP, Groothoff JW, Oosterveld FGJ. Self-reported functional status as predictor of observed functional capacity in subjects with early osteoarthritis of the hip and knee: a diagnostic study in the CHECK cohort. *J Occup Rehabil*. 2009;19(4):345–53.
19. Lamers LM, Meerding WJ, Severens JL, Brouwer WB. The relationship between productivity and health-related quality of life: an empirical exploration in persons with low back pain. *Qual Life Res*. 2005;14(3):805–13.
20. Ahlstrom L, Grimby-Ekman A, Hagberg M, Dellve L. The work ability index and single-item question: associations with sick leave, symptoms, and health—a prospective study of women on long-term sick leave. *Scand J Work Environ Health*. 2010;36(5):404–12.
21. Koopmanschap M, Meerding WJ, Evers S, Severens J, Burdorf A, Brouwer W. Productivity and disease questionnaire-PRO-DISQ. 2.1st ed. Rotterdam: Erasmus University; 2004.
22. Tait RC, Chibnall JT. Factor structure of the pain disability index in workers compensation claimants with low back injuries. *Arch Phys Med Rehabil*. 2005;86(6):1141–6.
23. Soer R, Reneman MF, Vroomen PC, Stegeman P, Coppes MH. Responsiveness and minimal clinically important change of the pain disability index in patients with chronic back pain. *Spine (Phila Pa 1976)*. 2012;37(8):711–5.
24. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. *Health Econ*. 1993;2(3):217–27.
25. Zee Kvd, Sanderman R. Het meten van de algemene gezondheidstoestand met de RAND-36, een handleiding. 2nd ed. Groningen, The Netherlands: RuG, research institute SHARE; 2012.
26. Vendrig A. De vragenlijst arbeidsreïntegratie. *Diagnostiek-Wijzer*. 2005;8:27–39.
27. Soer R, van der Schans CP, Geertzen JH, Groothoff JW, Brouwer S, Dijkstra PU, et al. Normative values for a functional capacity evaluation. *Arch Phys Med Rehabil*. 2009;90(10):1785–94.
28. Gross DP, Battie MC. Construct validity of a kinesiophysical functional capacity evaluation administered within a worker's compensation environment. *J Occup Rehabil*. 2003;13(4):287–95.
29. Keller A, Hellesnes J, Brox JI. Reliability of the isokinetic trunk extensor test, Biering-Sorensen test, and Astrand bicycle test: assessment of intraclass correlation coefficient and critical difference in patients with chronic low back pain and healthy individuals. *Spine (Phila Pa 1976)*. 2001;26(7):771–7.
30. Hodselmans AP, Dijkstra PU, Geertzen JH, van der Schans CP. Exercise capacity in non-specific chronic low back pain patients: a lean body mass-based Astrand bicycle test; reliability, validity and feasibility. *J Occup Rehabil*. 2008;18(3):282–9.
31. Bruce RA, Kusumi F, Hosmer D. Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. *Am Heart J*. 1973;85(4):546–62.
32. Bouwman C, De Jong K, Timman R, Hakkaart-Van Roijen L. Feasibility and validity of the TiC-P. *Value Health*. 2011;14(7):A426.
33. Dyer MT, Goldsmith KA, Sharples LS, Buxton MJ. A review of health utilities using the EQ-5D in studies of cardiovascular disease. *Health Qual Life Outcomes*. 2010;28(8):13.
34. Gronblad M, Hupli M, Wennerstrand P, Jarvinen E, Lukinmaa A, Kouri JP, et al. Intercorrelation and test-retest reliability of the pain disability index (PDI) and the Oswestry Disability Questionnaire (ODQ) and their correlation with pain intensity in low back pain patients. *Clin J Pain*. 1993;9(3):189–95.
35. Cleland JA, Childs JD, Whitman JM. Psychometric properties of the neck disability index and numeric pain rating scale in patients with mechanical neck pain. *Arch Phys Med Rehabil*. 2008;89(1):69–74.
36. Young IA, Cleland JA, Michener LA, Brown C. Reliability, construct validity, and responsiveness of the neck disability index, patient-specific functional scale, and numeric pain rating scale in patients with cervical radiculopathy. *Am J Phys Med Rehabil*. 2010;89(10):831–9.
37. Reneman MF, Dijkstra PU, Westmaas M, Goeken LN. Test-retest reliability of lifting and carrying in a 2-day functional capacity evaluation. *J Occup Rehabil*. 2002;12(4):269–75.
38. van Abbema R, Lakke SE, Reneman MF, van der Schans CP, van Haastert CJ, Geertzen JH, et al. Factors associated with functional capacity test results in patients with non-specific chronic low back pain: a systematic review. *J Occup Rehabil*. 2011;21(4):455–73.
39. Nordrehaug JE, Danielsen R, Stangeland L, Rosland GA, Vik-Mo H. Respiratory gas exchange during treadmill exercise testing: reproducibility and comparison of different exercise protocols. *Scand J Clin Lab Invest*. 1991;51(7):655–8.
40. Minnock P, Kirwan J, Bresnihan B. Fatigue is a reliable, sensitive and unique outcome measure in rheumatoid arthritis. *Rheumatology*. 2009;48:1533–6.