Content Comparison of Worker Productivity Questionnaires in Arthritis and Musculoskeletal Conditions Using the International Classification of Functioning, Disability, and Health Framework

Reuben Escorpizo · Alarcos Cieza · Dorcas Beaton · Annelies Boonen

Published online: 26 July 2009 © Springer Science+Business Media, LLC 2009

Abstract *Background* Worker productivity outcome is essential in examining the rehabilitation of workers with arthritis and other musculoskeletal conditions. There is great variation in the contents of worker productivity questionnaires. The International Classification of Functioning Disability and Health (ICF) offers the possibility to serve as a reference to describe and compare the contents of these questionnaires. *Methods* A literature review identified published self-report worker productivity questionnaires. All meaningful concepts were identified and linked to the corresponding ICF category according to

A. Cieza

ICF Research Branch of the World Health Organization-Collaborating Center, Family of International Classifications (DIMDI), Institute for Health and Rehabilitation Sciences, Ludwig-Maximilian University, Munich, Germany

D. Beaton

Mobility Program Clinical Research Unit, Keenan Research Centre, Li Ka Shing Knowledge Institute at St Michael's Hospital, Toronto, ON, Canada

D. Beaton

Institute for Work and Health, Toronto, ON, Canada

A. Boonen

Department of Internal Medicine, Division of Rheumatology, Maastricht University Medical Center, Maastricht, The Netherlands

A. Boonen Caphri Research Institute, Maastricht, The Netherlands established rules. Results Eighteen questionnaires were identified which contained a total of 519 meaningful concepts and which were linked to 64 unique 2nd level ICF categories. All questionnaires addressed Activities and Participation, thirteen (72%) addressed Body Functions, seven (39%) addressed Environmental Factors, seven (39%) addressed Personal Factors and only one questionnaire (6%) for Body Structures component. Overall, Work Role Functioning (WRF) questionnaire addressed the most number of different categories while Quantity and Quality method contained only one ICF category. The Rheumatoid Arthritis-Work Instability Scale had the highest number of categories for Body Functions, the Work Activity Limitations Scale and WRF had the most number of categories for Activities and Participation. The Health and Labour Questionnaire had the highest number of categories referring to unpaid work participation. The Health and Work Questionnaire was the only that included contextualization of both Environmental and Personal Factors. Conclusion Self-report worker productivity questionnaires differed largely in their contents. This content analysis study could guide us in selecting an appropriate questionnaire for a specific study question.

Keywords Outcomes assessment · Work · Questionnaires · Arthritis · Absenteeism · Presenteeism

Introduction

From a global perspective, musculoskeletal (MSK) conditions are a major contributor to physical disability, increased healthcare utilization, and reduced health-related quality of life [1]. In Europe, chronic pain and physical disability and high socioeconomic burden are associated

R. Escorpizo (⊠) · A. Cieza ICF Research Branch of the World Health Organization-Collaborating Center, Family of International Classifications, Swiss Paraplegic Research, Guido A. Zach Strasse 4, 6207 Nottwil, Switzerland e-mail: reubenescorpizo@aol.com

with MSK conditions, and this burden is expected to increase [2]. MSK conditions in Australia have high prevalence with notable consequences such as limited activity, reduced mobility, and increased risk for work disability—thus less productivity [3]. In a National Health Interview Survey in the United States, MSK conditions were the most prevalent compared to other self-reported medical problems in people 18 years and older, and with a cost-related burden of \$849 billion for 2002-2004 alone. Forty percent of this burden included work loss [4]. Arthritis, as an example of an MSK condition, affects the working age population, and has high personal and economic costs [5, 6]. Examining the impact of musculoskeletal and soft tissue disorders (including arthritis and chronic joint pain) on worker productivity has been of interest to rehabilitation therapists, economists, researchers in rehabilitation and outcome measurement, and patients. Several studies in the recent past have concentrated on withdrawal of patients from paid employment (i.e. sickness absence or sick leave) [7-11]. However, within the last decade, it has been recognized that patients also encounter problems while having paid employment. Hence, patients engaged in paid work can experience absence from work (termed absenteeism) and can also experience difficulties and reduced productivity while at work (termed presenteeism) [12]. Absenteeism (e.g. sick leave) and presenteeism are indicators of productivity at the level of the worker—referred to as *worker* productivity [13].

In order to measure worker productivity in rehabilitation and research studies, several self-report questionnaires have been developed. Self-report measures are useful particularly in examining worker productivity where there is no available objective measure (e.g. input-output units) [14]. The use of self-reporting has been shown in clinical and research settings, and could hold relevance from the public health and insurance perspectives [15-17]. However, self-report worker productivity questionnaires differ in their purpose and content [13]. Some questionnaires (or parts of them) are generic (being used regardless of the health condition) [18] and others are specific for a particular disease such as the Work Instability Scale for rheumatoid arthritis [19]. Some address only presenteeism like the Health and Work Questionnaire [20] or only absenteeism, or both like the Osterhaus Method [21] and the World Health Organization Health and Work Performance Questionnaire [14]. Some of questionnaires assess the effects of ill health on work while others also explore the impact of work on health. Some like the Work Limitations Questionnaire (WLQ) [22] includes a focus on output in order to quantitatively estimate monetary loss (e.g. number of hours), while others also explore qualitative experiences of the worker-work relationship (e.g. impact on life) [20]. Finally, there are also questionnaires that measure productivity in unpaid work in addition to paid work productivity [23].

The diversity of contents amongst questionnaires requires that we are able to judge which questionnaire is the most appropriate to answer a specific clinical research question. It is recognized that psychometric properties of validity, reliability, feasibility, and responsiveness are important in assessing questionnaires [24]. This study will focus on the content aspect of validity and the first step to address this issue would be to systematically identify and compare the concepts that are contained at the item level of these questionnaires using a reference framework.

Content examination would require an external reference framework that can be used as an interface between all questionnaires. The International Classification of Functioning, Disability, and Health (ICF) offers such a comparative interface. The ICF framework is based on the biopsychosocial model of functioning, disability, and health [25] and recognizes that functioning and health is a result of a complex interaction between and among its different components: Body Functions, Body Structures, Activities and Participation, Environmental Factors, and Personal Factors. Because of the broad and universal framework both at the individual level and societal level) of the ICF, it offers a comprehensive list of all categories necessary for the exhaustive description of functioning relevant to occupational rehabilitation. By identifying meaningful concepts that are conveyed and understood within the text of the items of the different worker productivity questionnaires and linking those concepts to the best fitting ICF category, it is possible to compare the content of these questionnaires.

In view of the broadness of the concepts addressed in the existing questionnaires that encompass functions, activities, participation issues as well as contextual factors, the ICF categories can be useful as a reference for comparison. Other similar ICF linking of measures in musculoskeletal conditions have been done [26–28] and proved useful in terms of creating a profile of how those measures mapped out well with the ICF. The current study would add evidence on how the contents of measures commonly used in occupational rehabilitation of MSK conditions and trials compare with the ICF—an area that has not been investigated until now. This study will apply MSK conditions in general to include arthritis as has been done in previous reports [1, 4] assuming that MSK conditions and symptoms.

The specific aims of the present study are to describe the content of self-report questionnaires that assess worker productivity and that are being used or could potentially be used in arthritis and other musculoskeletal conditions using the ICF as reference, and to compare contents of the questionnaires based on the linked concepts.

Materials and Methods

Questionnaires

The Worker Productivity Group of the Outcome Measures in Rheumatology (WPG-OMERACT) conducted a literature review focusing on patient-reported indicators of presenteeism as a function of worker productivity in arthritis and MSK conditions. Taken were measures from the review articles (published 2001-2004) and supplemented with questionnaires found in a separate literature review (up to 2002), and the grey literature. Search for key articles for each scale was conducted and citation searches on these articles were performed to locate any additional measures or any information on psychometric testing for any of the measures. Articles that assessed psychometric properties of validity, reliability, and responsiveness were included [13]. Developers were contacted for consent to use their questionnaires for this ICF linking project. Different versions of the questionnaires, if any, were included.

Questionnaire Characteristics

For each questionnaire, the general characteristics were noted, total number of items, number of items on sociodemographics of the respondents, and type of worker productivity that were addressed (absenteeism, presenteeism, unpaid work or combinations of these). Items related to the assessment of sociodemographics could be related to age, sex, race, educational level, income, marital status, occupation, employment status, time working, job content, type of work sector, size of company, other life commitments, and religion. Identification and Linking of Meaningful Concepts

To understand linking, it is important to have insight into the hierarchical system of the ICF classification. The ICF's hierarchical classification system contains 1,454 categories. A category is defined as the classification unit of a concept-that provides description to health and health-related domains of the entire ICF framework [25]. ICF categories are divided over five components: Body Functions, Body Structures, Activities and Participation and the contextual factors, comprising Environmental Factors and Personal Factors. At this time, Personal Factors component is not yet specified in categories. A category is assigned an identification code consisting of a letter referring to the ICF component it belongs to and a series of figures. The letter 'b' refers to Body Functions, 's' to Body Structures, 'd' to Activities and Participation, 'e' to Environmental Factors, and 'p' to Personal Factors. The category numbers refer to a hierarchical system referring the 2nd, 3rd, or 4th levels. The categories are grouped into chapters with a common theme. The 3rd and 4th level categories are specifications of the less specific higher level (i.e. 2nd level). An example of an ICF code is d850 for a 2nd level category for Remunerative Employment under the component Activities and Participation (d) and under chapter on Major Life Areas (chapter 8).

For each questionnaire, the content of the items was identified and linked to the ICF by two independent linkers (Fig. 1). Linking was performed following previously published linking rules [29, 30] and the major steps are outlined in Fig. 1.

In the qualitative part, meaningful concepts were identified in each questionnaire item. A meaningful concept is a "unit of text" identified to convey a single theme based on



the linker's judgement and expertise of functioning and the ICF. For example, the item "play cards and other games" contains the meaningful concept of "play" [29]. One questionnaire-item can include more than one meaningful concept. Non-items such as introduction, instructions, and item response options were not considered for linking. Only the actual worker productivity items of the questionnaire were linked, hence preliminary or supplementary sociodemographics questions regarding information (including those that may pertain to personal factors) were not considered. Although we decided not to link sociodemographics, it was to be documented. Also, those parts of questionnaires or items that were part of explanatory statement prior to the actual question or item were not linked. In the second part, meaningful concepts previously identified in the first part were compared and each concept was linked to an ICF category by identifying first the most appropriate chapter of the component and then the category of that chapter that most precisely maps to the meaning of the concept. If the content of a concept was more general than the corresponding ICF category, the code of the higher level was linked. Quality assurance was ascertained. The entire procedure was performed by two independent individuals (health professionals) who were trained in the linking process, have practical experience of the linking procedure, and who have had good level of agreement based on previous studies [27, 31]. In case of disagreement between both linkers, a third person trained in linking rules was consulted and made an informed decision (AC).

To evaluate the reliability of the linking process, the overall percentage of agreement for each questionnaire was calculated based on the two independent linkage versions. In addition, to examine the extent to which the achieved agreement exceeds chance, the Kappa coefficient [32] and nonparametric bootstrapped confidence interval [33, 34] were calculated. We performed Kappa analysis using SAS software version 9.1 (Copyright[©] 2002–2003 by SAS Institute Inc., Cary, NC, USA).

As an illustration of the process, let us take item 9 of the Rheumatoid Arthritis-Work Instability Scale (RA-WIS) "*I am very worried about my ability to keep working*", will be linked to the ICF. In this item, two meaningful concepts were identified which could be linked to two different ICF categories:

Concept 1: Worrying linked to b160 (thought functions) Concept 2: To keep working linked to d850 (remunerative employment)

As part of the linking rules it is important to realize that a meaningful concept which refers to a *diagnosis or a health condition* such as "symptoms of your arthritis" is not linked and was labelled "health condition ("hc"). If a concept related to a *Personal Factor*, it was labelled personal factor ("pf"). A concept that was not contained in the ICF universe of categories was labelled "not covered" ("nc"). Finally, if the information about a concept was not sufficient to make a decision about the most appropriate ICF category, it was labelled "not definable" ("nd"). "Nc" and "nd" could be specified further to belong to specific ICF components (e.g. nd-*Body Functions*, nd-*Activities and Participation*, etc.). "Nc" and "nd" concepts were pooled for the purpose of our analysis and were labelled collectively as "not covered".

For each questionnaire, the number of meaningful concepts, the number of ICF categories (total and 2nd level) and their distribution over the ICF components and types of ICF categories overall and per component were reported to make a basis for comparing all the questionnaires.

Results

Questionnaires

The review strategy of WPG-OMERACT revealed 18 selfreport questionnaires which are presented in Table 1. There were seven questionnaires from the review articles, five

 Table 1
 Overview of questionnaires

Questionnaires	Number	Domains cove	ered by the iter	ns
	of items	Socio- demographic items	Absenteeism	Presenteeism
EWPS	25	Yes	Yes	Yes
HLQ	17	Yes	Yes	Yes
HPQ	13	Yes	Yes	Yes
HRPQ-D	9	No	Yes	Yes
HWQ	24	No	No	Yes
LFQ	16	Yes	Yes	Yes
OST	12	Yes	Yes	Yes
QQ	2	No	No	Yes
RA-WIS ^a	23	No	No	Yes
SPS-6	6	No	No	Yes
SPS-13	13	No	Yes	Yes
WALS ^a	12	No	No	Yes
WLQ-25	25	No	No	Yes
WLQ-16	16	No	No	Yes
WLQ-8	8	No	No	Yes
WPAI-GH	6	No	Yes	Yes
WPSI ^b	9	Yes	Yes	Yes
WRF	26	No	No	Yes

^a Arthritis-specific questionnaire

^b Arthritis-specific module within questionnaire

questionnaires from the separate literature review, and two questionnaires from the grey literature. We then identified 24 key articles for each scale and conducted citation searches on these articles to locate any additional measures or any information on psychometric testing for any of the measures. This citation search yielded 198 articles, from which we gathered 2 more questionnaires [13]. With WLQ 16- and 8-item versions treated separately, there were 18 questionnaires in total. The questionnaires include the following: Endicott Work Productivity Scale (EWPS) [35], Health and Labour Questionnaire (HLQ) [36], WHO Health and Work Performance Questionnaire-Clinical Trials Baseline version (HPQ) [14], Health-Related Productivity Questionnaire Diary (HRPQ-D-D) [37], Health and Work Questionnaire (HWQ) [20], Life Functioning Questionnaire (LFQ) [23], Osterhaus technique (OST) [21], Quantity and Quality Method (QQ) [38], Rheumatoid Arthritis-Work Instability Scale (RA-WIS) [19], Stanford Presenteeism Scale-6 items (SPS-6) [39], Stanford Presenteeism Scale-13 items (SPS-13) [40], Work Activity Limitations Scale (WALS) [41], Work Limitations Questionnaire-25 items (WLQ-25) [22], Work Limitations Questionnaire-16 items (WLQ-16) [42], Work Limitations Questionnaire-8 items (WLQ-8) [43], Work Productivity and Activity Impairment-General Health (WPAI-GH) [18], Work Productivity Short Inventory-arthritis section (WPSI) [44], and Work Role Functioning/Work Limitations-26 items (WRF) [45]. HRPQ-D was referred to as the shorter diary format of the HRPQ-D (Steve Hass, personal communication). The 18 measures presented in this article included four versions of WLQ (WLQ-8, WLQ-16, WLQ-25 and WRF) and two versions of SPS (SPS-6 and SPS-13). RA-WIS and WALS were arthritis-specific questionnaires while WPSI only has a section on arthritis and rheumatism.

Table 1 shows the number of items included in each questionnaire and indicates the different domains covered. Six questionnaires contained items (sometimes as a separate module) on sociodemographics, all questionnaires included items that addressed presenteeism and nine questionnaires covered both presenteeism and absenteeism from paid work.

Linking and ICF Coding

The 18 questionnaires contained 262 items. Seventeen items were not considered for linking based on the exclusion criteria above (e.g. sociodemographics, instructions, etc.). Hence, 245 items remained. From the remaining 245 items, 519 meaningful concepts were extracted and linked. On average, a single item contained 2.3 meaningful concepts. The 519 concepts were linked to 64 unique 2nd level ICF categories. From these 64 categories, 13 categories

(21%) belonged to *Body Functions*, one category (2%) to *Body Structures*, 40 (63%) to *Activities and Participation* and 10 (16%) to *Environmental Factors*). Fourteen concepts referred to *Personal Factors* but not categorized according to current ICF version and 20 meaningful concepts could not be linked to the ICF because they were not covered (nc) or not defined (nd).

Table 2 present the number of items linked, number of meaningful concepts identified, distribution of all as well as 2nd level linked categories over ICF components. In addition, Table 2 presents the number of *Personal Factors* and the number of concepts that could not be linked (nd + nc).

It can be seen that no questionnaire addressed all ICF components. All questionnaires contained categories belonging to the component *Activities and Participation*, thirteen (72%) contained categories belonging to *Body Functions*, seven (39%) belonging to *Environmental Factors*, and one questionnaire (6%) belonging to *Body Structures* component. Seven questionnaires (39%) included concepts that were *Personal Factors*. The WRF addressed the highest number of 2nd level categories (n = 17). QQ, with only two items, addressed only one category (d850 *Remunerative Employment*).

Detailed results of linking of the concepts to specific ICF components and 2nd level categories corresponding to the ICF contents of the questionnaires are provided in Table 3.

All questionnaires mainly covered categories belonging to the component Activities and Participation. Within the component Activities and Participation, the category most frequently addressed was d850 Remunerative Employment. This category was addressed by all questionnaires. Other categories from this component that were frequently addressed were General Tasks and Demands (d210-d240), Carrying, Moving and Handling Objects (d430-d445), and Interpersonal Relationships (d710-d760). Less frequently addressed were d820-d839 under Education (LFQ, HRPQ-D, OST, and WPAI-GH); and Acquisition of Necessities (d620), Household Tasks (d630-d640) or Caring for Household Objects and Assisting Others (d650-d660) (LFQ, HLQ, HRPQ-D, and WPAI-GH), and Recreation and Leisure (d920) (HLQ, HPQ, and WPAI-GH (e.g. time missed from work because of leisure, vacation, and time-off)).

The WALS and WRF had the highest number of categories for the component *Activities and Participation* (n = 14). HLQ covered the highest number of concepts on non-paid work participation (n = 10) such as *Acquisition of Necessities* (like d6200 *Shopping*), *Household Tasks* (like d640 *Housework*), and *Caring for Household Objects* (d650).

The next most frequently addressed component was *Body Function*. Within this component, the categories most

Number of items linked 25 16 (<i>n</i>) Number of concepts (<i>n</i>) 42 55 Number of ICF categories per component Body functions 7 (5) 5 (4) [total (2nd level)]	12	нүү нкүчл-г	DWH (OST (4 S	SA-WIS S	3-S-6	SPS-13	WALS	WLQ-25	WLQ-16	WLQ-8	WPAI-GH	ISdM	WRF
Number of concepts (n) 4255Number of ICF categories per component Body functions7 (5)5 (4)[total (2nd level)]7 (5)5 (4)	15	6 6	24	∞	5	23 6		ε.	12	25	16	8	9	5	26
Number of ICF categories per component Body functions 7 (5) 5 (4) [total (2nd level)]	f	42 21	46	6	5	38 1	1	ŝ	32	43	21	6	23	16	41
Body functions 7 (5) 5 (4) [total (2nd level)]															
	1 (1)	1 (1) 0 (0)	5 (4)	0 (0) 0	3 (0) C	3 (7) 2	(2)	t (4)	1 (1)	4 (4)	3 (3)	1 (1)	(0) 0	3 (3)	3 (3)
Body structures 0 (0) 0 (0) [total (2nd level)]	(0) 0	0 (0) 0 (0)	0 (0)	0 (0) 0) (0) C	0 (0) ((0)	(0)	(0) 0	0 (0)	(0) 0	(0) 0	(0) 0	1 (1)	(0) 0
Activities and 9 (6) 13 (10) participation [total (2nd level)]	18 (10)	3 (2) 5 (4)	6) (6)	2 (2)	1 (1) 3	3 (3) 2	(2)	(4)	21 (14)	16 (11)	11 (8)	1 (1)	(L) L	1 (1)	22 (14)
Environmental 1 (1) 1 (0) factors [total (2nd level)]	4 (3)	1 (1) 0 (0)	4 (4)) (0) 0) (0) (0 (0) ((0)	(0)	(0) 0	(0) 0	(0) 0	1 (1)	0 (0)	(0) 0	(0) 0
Personal factors 0 0	0	0 0	8) 0	0 1	-	0	_	0	1	1	1	0	0	1
"Nc" + "nd" 0 3 concepts (n)	5	7 0	0	0	0 1	0	0	-	0	1	1	1	3	0	1
Number of different 17 (12) 19 (14) ICF categories [total (2nd level)]) 23 (14)	5 (4) 5 (4)	18 (14)	2 (2)	1 (1) 1	11 (10) 4	t (4) 5	(8)	22 (15)	20 (15)	14 (11)	3 (3)	(L) L	5 (5)	25 (17)

The table shows the total number of unique categ presented. See text for questionnaire acronyms



Table 3 continued																				
ICF Category	EWPS	LFQ	HLQ	ЫРQ	HRP(Q-D F) DWI	OST 4	QQ R	SIM-A.	SPS-6	SPS13	WAL	S WLC	225 WI	_Q16	WLQ8	WPAI-GH	ISdM	WRF
d175 Solving problems																				
d177 Making decisions																				
Chapter 2: General tasks and demands																				
d210 Undertaking a single task													-							-
d220 Undertaking multiple tasks	•					-	_					-			-					-
d230 Carrying out daily routine	•																	•		
d240 Handling stress and other psychological										_										
demands																				
Chapter 3: Communication						-	_							•						
d360 Using communication																				
devices and techniques													•	•						
Chapter 4: Mobility																				
d410 Changing basic body position													•	•						-
d415 Maintaining a body position													•		•					-
d430 Lifting and carrying objects													•	•	•					
d435 Moving objects with lower extremities																				-
d440 Fine hand use										_										
d445 Hand and arm use													-	•						
d450 Walking																				
d455 Moving around																				
d460 Moving around in different locations																				
d470 Using transportation																				
d475 Driving																				
Chapter 6: Domestic life																				
d620 Acquisition of goods and services			-																	
d630 Preparing meals																				





frequently addressed were related to *Specific Mental Functions*, (such as b140 *Attention Functions* and b152 *Emotional Functions*) and *Global Mental Function* (such as b1300 *Energy Level* and b1301 *Motivation*). Chapters on *Sensory Functions and Pain, Functions of the Joints and Bones* (i.e. *Mobility of Joints*) and *Movement Functions* were only addressed in RA-WIS and WPSI. RA-WIS contained the highest number of 2nd level categories (n = 7) under *Body Functions*.

The component *Body Structures* was only contained within the WPSI (s7701 *Joints*) and was a specification of the category *Additional Musculoskeletal Structures Related* to Movement.

Only few questionnaires (EWPS, LFQ, HLQ, HPQ, HWQ, and WLQ-8) addressed the component *Environmental Factors* and only a limited number of chapters and categories were addressed. The ten environmental categories covered *Assistive Products and Technology for Employment* (e1351), *Sound Intensity* (e2500), *Indoor Air Quality* (e2600) and six categories belonged to the chapters *Support and Relationships* (e3) and *Attitudes* (e4). Except for HWQ, *Environmental Factors* were specifications to describe the context of hindrance at work.

The questionnaires that included concepts belonging to *Personal Factors* were HWQ, RA-WIS, SPS-6 and all versions of WLQ (including WRF). In the majority of questionnaires, the *Personal Factors* referred to the 'level of satisfaction' regarding work (HWQ and SPS-6) or work performance (WLQ-8, 16, 25 and WRF). HWQ addressed also the "level of control" over aspects of work. Also, the concept contained in the item "I feel I have to give up my work" of RA-WIS was considered a personal factor indicating coping style or behaviour relevant to the experience and life events of the worker and is not necessarily part of his or her health state (as defined in the ICF) [25] HWQ had the highest number of combined *Personal and Environmental Factors*.

Within the nine questionnaires (LFQ, HLQ, HPQ, RA-WIS, WLQ-25, WLQ-16, WLQ-8, WPAI-GH, and WRF), a total of 20 meaningful concepts were identified that could not be linked to ICF categories- 14 of the 20 were not sufficiently defined (nd) to be linked to a specific ICF category and the remaining six were not covered (nc) by the ICF. "Nd" concepts mainly included specifications of physical or mental functioning, role and health, or general description of functioning.

Table 4 shows the results of the evaluation of the linkage procedure by the overall percentage of agreement, Kappa statistics, and bootstrapped confidence intervals. The overall percentage of agreement ranges from 38.2 to 100% and the estimated Kappa values from 0.35 to 0.8. None of the 95% confidence intervals encloses zero, thus the level of agreement exceeded chance.
 Table 4
 Overall percentage of agreement, Kappa coefficient, and nonparametric bootstrapped confidence interval for each of the questionnaires studied

Questionnaires	Overall percentage of agreement (%)	Kappa coefficient	Nonparametric bootstrapped confidence interval
EWPS	67.0	0.44	(0.29, 0.64)
HLQ	38.2	0.35	(0.23, 0.51)
HPQ	75.0	0.67	(0.40, 1.00)
HRPQ-D	77.3	0.77	(0.59, 1.00)
HWQ	59.1	0.65	(0.54, 0.79)
LFQ	47.1	0.63	(0.53, 0.74)
Osterhaus	100	_	_
QQ	100	_	_
RA-WIS	76.3	0.49	(0.30, 0.75)
SPS-6	56.3	0.40	(0.01, 0.89)
SPS-13	55.6	0.80	(0.71, 0.94)
WALS	68.8	0.75	(0.61, 0.94)
WLQ-25	68.9	0.73	(0.66, 0.84)
WLQ-16	73.7	0.68	(0.53, 1.00)
WLQ-8	70.0	0.47	(0.15, 1.00)
WPAI-GH	75.0	0.56	(0.40, 1.00)
WPSI	68.9	0.54	(0.40, 0.69)
WRF	66.7	0.66	(0.47, 0.94)

Discussion

This study analyzed and compared the contents of 18 selfreport worker productivity questionnaires that can be used in arthritis and MSK conditions. The analysis was based on the ICF and provided relevant information on the content of these questionnaires. The results can guide researchers in selecting and reporting worker productivity questionnaires in clinical trials or longitudinal studies in which worker productivity is among the endpoints and also which questionnaires could be used by rehabilitation professionals in the clinical practice. Specifically, selection could be made depending on which of the questionnaires contain aspects of worker productivity (using the ICF as the basis) that would be relevant to the aim of a particular study. ICFbased contents validity as shown in this study could guide such selection in addition to questionnaire reliability, construct validity, feasibility, and responsiveness.

This study does not aim to propose which of the questionnaires are preferred. The questionnaires have been developed for different purposes and their approach to worker productivity varies, reflecting the complexity of work as an outcome measure. Not only is worker productivity a multi-domain outcome (absenteeism, withdrawal, sick leave, presenteeism, etc.) but it also has varying perspectives from different stakeholders (patients, employers, managers, social security systems, society, etc.), and hence may hold different relevance for different settings such as the type of jobs and types of industries. Qualitative aspects of the work-worker relationship might be important for patients but could be less important for the employer and in business decision-making. A wide variation in questionnaires was evident in our findings-differences in the conceptual definition and operationalization of worker productivity (absenteeism, presenteeism or both; amount of time, level of difficulty, or degree of impairment), focused versus broad approach to capturing information, and a mix of scales and subscales. Nevertheless, all questionnaires addressed the category d850 remunerative employment, which was not surprising considering their purpose of examining worker productivity. Limitations in specific activities can provide invaluable insights into restrictions in worker productivity or could be a specific target of treatment, among which is occupational rehabilitation programs.

Unpaid work has been increasingly recognized to be also important in the area of worker productivity in general and its importance has been reflected in some of the questionnaires. Some questionnaires also addressed restrictions in areas of participation other than paid work. LFQ, HLQ, HRPQ-D, and WPAI-GH addressed unpaid work such as *Household tasks*. *Education* (as a student or one that engages in academic activities) was addressed by four questionnaires (LFQ, HRPQ-D, OST, and WPAI-GH). *Recreation and leisure* in the form of play, sports, hobbies, and crafts were addressed by three questionnaires (HLQ, HPQ, and WPAI-GH). Interestingly, engagement in sports activities for example, was found to influence work

Table 5 Comparison of the 2nd-level contents of examined questionnaires with the ICF core set on arthritis and MSK clinical trials

ICF categories contained in questionnaires	ICF core set for RA ^a	ICF core set for OA ^b	MSK clinical trials
b130 Energy and drive function			
b134 Sleep functions	•		
b140 Attention functions			
b144 Memory functions			
b152 Emotional functions			•
b160 Thought function			•
b210 Seeing functions			•
b270 Sensory functions related to temperature and other stimuli			•
b280 Sensation of pain	•		•
b710 Mobility of joint functions	•		•
b780 Sensation of muscle stiffness	•		•
s770 Additional musculoskeletal structures related to movement	•		•
d163 Thinking			
d166 Reading			
d170 Writing	•		
d175 Solving problems			
d177 Making decisions			
d210 Undertaking a single task			
d220 Undertaking multiple tasks			
d230 Carrying out daily routine	•		
d240 Handling stress and other psychological demands			
d350 Conversation			
d360 Using communication devices and techniques			
d410 Changing basic body positions			
d415 Maintaining a body position			
d430 Lifting and carrying objects			
d435 Moving objects with lower extremities			
d440 Fine hand use			•
d445 Hand and arm use			
d450 Walking			•
d455 Moving around			
d460 Moving around in different locations			•
d470 Using transportation			•
d475 Driving	•		•
d620 Acquisition of goods and services			•
d630 Preparing meals	•		•
d640 Doing housework			•
d650 Caring for household objects			•
d660 Assisting others	•		
d710 Basic interpersonal interactions			
d720 Complex interpersonal interactions			
d740 Formal relationships			
d750 Informal social relationships			
d760 Family relationships			
d820 School education			
d839 Education, other specified and unspecified			
d845 Acquiring, keeping, and terminating a job			
d850 Remunerative employment		•	•

Table 5 continued

ICF categories contained in questionnaires	ICF core set for RA ^a	ICF core set for OA ^b	MSK clinical trials
d860 Basic economic transactions			
d920 Recreation and leisure			
e135 Products and technology for employment			
e250 Sound			
e260 Air quality			
e320 Friends			
e325 Acquaintances, peers, colleagues, neighbours and community members			
e340 Personal care providers and personal assistants			
e425 Individual attitudes of acquaintances, peers, colleagues, neighbours and community members	•		
e430 Individual attitudes of people in positions of authority			

■ Indicates a match

^a Rheumatoid arthritis

^b Osteoarthritis

productivity [46]. Therefore, it might be worthwhile to assess at the same time restrictions in paid as well as unpaid areas of participation in clinical evaluation or intervention study so as to gain insight into the relationships between these two areas and address occupational rehabilitation properly from within and beyond the confines of the workplace.

Arthritis- or MSK-relevant *Body functions* such as *Sensitivity to pressure, Pain in joints, Joint mobility,* and *Sensation of muscle stiffness* were covered by RA-WIS, and WPSI. This observation was not unexpected because these are arthritis-specific questionnaires or have a part that is relevant to arthritis (i.e. WPSI). Information gained from these *Body function* categories is important in further understanding the reason around limited worker participation specifically in patients with arthritis or MSK conditions.

Although contextual factors (*Environmental and Personal Factors*) are essential when exploring and interpreting the relationships between health and work, there were only a few questionnaires that addressed *Environmental factors* and *Personal factors*. Environmental categories were usually pertinent to attitudes and support from colleagues and managers and *Personal Factors* usually addressed satisfaction with work or work performance or sense of control over work. Low job satisfaction has been associated with productivity loss among workers with musculoskeletal conditions [47] and a questionnaire that could capture this information amongst other "personal attributes" would be helpful in further defining worker productivity.

When discussing contextual factors, it should be realized that several questionnaires separately assessed sociodemographic factors such age, gender, type of job, hours on the job, and income and as such may personal factors but were not included for linking since many of such demographic factors are usually part of questionnaires and often need to be adapted to the socioeconomic situation of each country anyway. Nevertheless, it needs to be emphasized that the inclusion of these factors being an aspect of work is essential in any study examining worker productivity.

With respect to the linking methodology, the linkage process was evaluated by calculating the overall percentage of agreement and Kappa coefficients with their confidence intervals. Based on the findings, there was not only satisfactory agreement between the linkers but also that the agreement exceeded chance. This shows that the linking procedure was reliable in this study.

Linking was proven to be useful not only to gain insight into the concepts of the questionnaires but also revealed other aspects of the contents of the questionnaires. For example, it was often seen that one item contained several concepts which can result in a difficulty to interpret as to which of the different concept the answer pertains to. An example is the WPSI item "... *pain, swelling, stiffness and loss of function in joints*" as factors that might affect work. Including more concepts in one item can help to reduce the number of items (feasibility) but also risks of losing specificity when interpreting the scores. In other words, on one hand, the brevity of questionnaires can reduce administration burden but on the other hand may sacrifice quantification of multiple concepts embedded in a single question.

It also became also clear that there is no absolute relationship between the number of items included for linking and the number of categories (i.e. 2nd level) addressed in the items. Not only can one item contain several categories, but also different items can address only one category. The HPQ, for example, has 13 items but only addressed 4 categories (Attention functions, Remunerative employment, Recreation and leisure, and Individual attitudes of people in position of authority). Items that address the same category can provide essentially different information. Some examples of items of the HPQ that were linked to the category d850 Remunerative employment were 'did you miss an entire workday', 'did you miss part of a workday', 'was your performance at work reduced' or 'was the quality of your work lower'. This indicates the multiple characterizations of the amount and type of restrictions that workers can encounter given a single category on remunerative work.

The findings of this study can also be applied to existing ICF-related studies such as those in content examination and ICF core sets. As presented in Table 5, 53% (31 of 58 2nd level) categories identified in the worker productivity questionnaires were also included in the ICF core set for osteoarthritis [48] and rheumatoid arthritis [49], while 55% (32 of 58) were found in measures used in MSK trials [26]. This highlights that the arthritis core sets and measures used in MSK trials represent at least half of the worker productivity questionnaires we examined contents-wise. Some categories we found did not match any of the core sets' such as d435 Moving objects with lower extremities, e250 Sound, e260 Air quality, and e430 Individual attitudes of people in positions of authority which we believe are specific to work or within a specific employment setting or specific condition. So, perhaps along with the chosen questionnaire, a core set can be used as a separate module in worker productivity studies that look at workers with arthritis or MSK conditions.

Several limitations were encountered in our analyses. Firstly, the linking process as a method does not consider the contextualization that is included in an item. For example, in the WRF item "...bend, twist, or reach while working..." the linking procedure linked four concepts "bending", "twisting", "reaching", and "work". However, it can be stated that "work" in this sense contextualizes all the activities of bending, twisting, and reaching and this contextualization is not captured by the linking procedure. Secondly, linking does not include time recall included in the item nor does it consider the directionality of cause and effect, if any. The directionality of concepts within one item is lost after identification of relevant ICF categories. For example, the item "... stress of my job makes my arthritis flare" from RA-WIS refers to the adverse influence of a work characteristic on health. On the other hand, the item "How much difficulty do you have working with your hands" from WALS ask about the impact of impairments in the hands on work limitation. Thirdly, while quality of linking was assured, the identification of meaningful concepts and coding based on the ICF of the two linkers not only require an in depth understanding of the ICF but also experience in linking. Hence, the degree of dissent between linkers should always be considered. Fourthly, sociodemographics information such as age, gender, profession, and education could be considered personal factors. This study did not consider sociodemographics, it only focused on the actual worker productivity items of the questionnaires. Finally, several meaningful concepts were not covered or not defined within ICF which might have resulted in loss of real information from the questionnaires. In RA-WIS, the linking team considered the concept included in the item "I used my holiday so that I don't have to go sick" as not covered and this was also the case for the concept "satisfying those people who judge your work" from the WLQ16 and WRF. It could be argued that these concepts are Personal factors, referring to an active coping style and an individual's effort of pleasant work behaviour in RA-WIS and WLQ16/WRF, respectively. In HPQ, the concepts in the item "...rate the usual performance of most workers in a job similar to yours?" and "...you were a lot better than other workers..." were considered "not covered" although these items clearly aim to quantify worker productivity (Remunerative Employment) using a reference or anchor point.

Work is a major life area to most people and is an important outcome or endpoint to measure in clinical studies and evaluation. The use of self-report questionnaires is important in advancing healthcare management around an individual's work participation and disability. Using the ICF as an external template to compare the contents of work productivity questionnaires not only revealed that a broad range of ICF categories were assessed by these questionnaires but also confirmed large differences in the contents. Knowledge of the contents of a questionnaire can help to capture the specific information regarding worker productivity that researchers and clinicians want to examine.

Acknowledgments We would like to thank the ICF Research Branch for the kind support, Edda Amman who assisted in the linking process, and Barbara Kollerits who provided statistical support and also assisted in linking. Special thanks to OMERACT, collaborators of this study, questionnaire developers and their support staff, and Brittany Norton for providing assistance in preparing the manuscript.

References

1. World Health Organization. The burden of musculoskeletal conditions at the start of the new millenium. WHO Technical Report Series 919; 2002.

- Woolf AD, Akesson K. Can we reduce the burden of musculoskeletal conditions? The European action towards better musculoskeletal health. Best Pract Res Clin Rheumatol. 2007;21:1–3.
- Australian Institute of Health and Welfare 2005. Arthritis and musculoskeletal conditions in Australia. AIHW Cat. No. PHE67; 2005.
- United States Joint and Bone Decade. The burden of musculoskeletal diseases in the United States. Rosemont: American Academy of Orthopedic Surgeons; 2008. p. 1–19.
- Maetzel A, Li LC, Pencharz J, Tomlinson G, Bombardier C, Community Hypertension and Arthritis Project Study Team. The economic burden associated with osteoarthritis, rheumatoid arthritis, and hypertension: a comparative study. Ann Rheum Dis. 2004;63:395–401.
- 6. Badley EM, Wang PP. The contribution of arthritis and arthritis disability to nonparticipation in the labor force: a Canadian example. J Rheumatol. 2001;28:1077–82.
- Zirkzee EJ, Sneep AC, de Buck PD, Allaart CF, Peeters AJ, Ronday HK, et al. Sick leave and work disability in patients with early arthritis. Clin Rheumatol. 2008;27:11–9.
- Upmark M, Borg K, Alexanderson K. Gender differences in experiencing negative encounters with healthcare: a study of long-term sickness absentees. Scand J Public Health. 2007;35: 577–84.
- van der Klink JJ, Blonk RW, Schene AH, van Dijk FJ. Reducing long term sickness absence by an activating intervention in adjustment disorders: a cluster randomised controlled design. Occup Environ Med. 2003;60:429–37.
- Bultmann U, Sherson D, Olsen J, Hansen CL, Lund T, Kilsgaard J. Coordinated and tailored work rehabilitation: a randomized controlled trial with economic evaluation undertaken with workers on sick leave due to musculoskeletal disorders. J Occup Rehabil. 2009;19:81–93.
- Bogefeldt J, Grunnesjo MI, Svardsudd K, Blomberg S. Sick leave reductions from a comprehensive manual therapy programme for low back pain: the Gotland low back pain study. Clin Rehabil. 2008;22:529–41.
- Boles M, Pelletier B, Lynch W. The relationship between health risks and work productivity. J Occup Environ Med. 2004;46:737– 45.
- Escorpizo R, Bombardier C, Boonen A, Hazes JM, Lacaille D, Strand V, et al. Worker productivity outcome measures in arthritis. J Rheumatol. 2007;34:1372–80.
- Kessler RC, Barber C, Beck A, Berglund P, Cleary PD, McKenas D, et al. The World Health Organization Health and Work Performance Questionnaire (HPQ). J Occup Environ Med. 2003;45: 156–74.
- Loeppke R, Taitel M, Richling D, Parry T, Kessler RC, Hymel P, et al. Health and productivity as a business strategy. J Occup Environ Med. 2007;49:712–21.
- Roelen CA, van der Pol TR, Koopmans PC, Groothoff JW. Identifying workers at risk of sickness absence by questionnaire. Occup Med (Lond). 2006;56:442–6.
- Wang PS, Simon GE, Avorn J, Azocar F, Ludman EJ, McCulloch J, et al. Telephone screening, outreach, and care management for depressed workers and impact on clinical and work productivity outcomes: a randomized controlled trial. JAMA. 2007; 298:1401–11.
- Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics. 1993;4:353–65.
- Gilworth G, Chamberlain MA, Harvey A, Woodhouse A, Smith J, Smyth MG, et al. Development of a work instability scale for rheumatoid arthritis. Arthritis Rheum. 2003;49:349–54.
- 20. Shikiar R, Halpern MT, Rentz AM, Khan ZM. Development of the Health and Work Questionnaire (HWQ): an instrument for

assessing workplace productivity in relation to worker health. Work. 2004;22:219–29.

- Osterhaus JT, Gutterman DL, Plachetka JR. Healthcare resource and lost labour costs of migraine headache in the US. Pharmacoeconomics. 1992;2:67–76.
- Lerner D, Amick BC 3rd, Rogers WH, Malspeis S, Bungay K, Cynn D. The work limitations questionnaire. Med Care. 2001;39: 72–85.
- Altshuler L, Mintz J, Leight K. The life functioning questionnaire (LFQ): a brief, gender-neutral scale assessing functional outcome. Psychiatry Res. 2002;112:161–82.
- Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for outcome measures in rheumatology. J Rheumatol. 1998;25: 198–9.
- 25. World Health Organization. International classification of functioning, disability, and health: ICF. Geneva: WHO; 2001.
- 26. Brockow T, Cieza A, Kuhlow H, Sigl T, Franke T, Harder M, et al. Identifying the concepts contained in outcome measures of clinical trials on musculoskeletal disorders and chronic widespread pain using the international classification of functioning, disability and health as a reference. J Rehabil Med. 2004;44(Suppl):30–6.
- 27. Sigl T, Cieza A, Brockow T, Chatterji S, Kostanjsek N, Stucki G. Content comparison of low back pain-specific measures based on the international classification of functioning, disability and health (ICF). Clin J Pain. 2006;22:147–53.
- Stamm T, Geyh S, Cieza A, Machold K, Kollerits B, Kloppenburg M, et al. Measuring functioning in patients with hand osteoarthritis—content comparison of questionnaires based on the international classification of functioning, disability and health (ICF). Rheumatology (Oxford). 2006;45:1534–41.
- Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun B, Stucki G. ICF linking rules: an update based on lessons learned. J Rehabil Med. 2005;37:212–8.
- Cieza A, Brockow T, Ewert T, Amman E, Kollerits B, Chatterji S, et al. Linking health-status measurements to the international classification of functioning, disability and health. J Rehabil Med. 2002;34:205–10.
- 31. Geyh S, Kurt T, Brockow T, Cieza A, Ewert T, Omar Z, et al. Identifying the concepts contained in outcome measures of clinical trials on stroke using the international classification of functioning, disability and health as a reference. J Rehabil Med. 2004;44(Suppl):56–62.
- Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37–46.
- Efron B. The jackknife, the bootstrap and other resampling plans. Philadelphia: Society for industrial and applied mathematics; 1982.
- Vierkant RA. A SAS macro for calculating bootstrapped confidence intervals about a Kappa coefficient [homepage on the Internet]. Available from: http://www2.sas.com/proceedings/ sugi22/STATS/PAPER295.PDF.
- Endicott J, Nee J. Endicott work productivity scale (EWPS): a new measure to assess treatment effects. Psychopharmacol Bull. 1997;33:13–6.
- 36. van Roijen L, Essink-Bot ML, Koopmanschap MA, Bonsel G, Rutten FF. Labor and health status in economic evaluation of health care. The health and labor questionnaire. Int J Technol Assess Health Care. 1996;12:405–15.
- 37. Kumar RN, Hass SL, Li JZ, Nickens DJ, Daenzer CL, Wathen LK. Validation of the health-related productivity questionnaire diary (HRPQ-D) on a sample of patients with infectious mono-nucleosis: results from a phase 1 multicenter clinical trial. J Occup Environ Med. 2003;45:899–907.
- Koopmanschap M, Meerding WJ, Evers S, Severens J, Burdorf A, Brouwer W. Productivity and disease questionnaire-PRODISQ versie 2.1 (in Dutch). 2004.

- Koopman C, Pelletier KR, Murray JF, Sharda CE, Berger ML, Turpin RS, et al. Stanford presenteeism scale: health status and employee productivity. J Occup Environ Med. 2002;44:14–20.
- Turpin RS, Ozminkowski RJ, Sharda CE, Collins JJ, Berger ML, Billotti GM, et al. Reliability and validity of the Stanford presenteeism scale. J Occup Environ Med. 2004;46:1123–33.
- Gignac MA, Badley EM, Lacaille D, Cott CC, Adam P, Anis AH. Managing arthritis and employment: making arthritis-related work changes as a means of adaptation. Arthritis Rheum. 2004; 51:909–16.
- Beaton DE, Kennedy CA. Beyond return to work: testing a measure of at-work disability in workers with musculoskeletal pain. Qual Life Res. 2005;14:1869–79.
- Burton WN, Chen CY, Conti DJ, Pransky G, Edington DW. Caregiving for ill dependents and its association with employee health risks and productivity. J Occup Environ Med. 2004;46: 1048–56.
- 44. Goetzel RZ, Ozminkowski RJ, Long SR. Development and reliability analysis of the work productivity short inventory (WPSI)

instrument measuring employee health and productivity. J Occup Environ Med. 2003;45:743–62.

- 45. Amick BC 3rd, Lerner D, Rogers WH, Rooney T, Katz JN. A review of health-related work outcome measures and their uses, and recommended measures. Spine. 2000;25:3152–60.
- 46. van den Heuvel SG, Heinrich J, Jans MP, van der Beek AJ, Bongers PM. The effect of physical activity in leisure time on neck and upper limb symptoms. Prev Med. 2005;41:260–7.
- van den Heuvel SG, Ijmker S, Blatter BM, de Korte EM. Loss of productivity due to neck/shoulder symptoms and hand/arm symptoms: results from the PROMO-study. J Occup Rehabil. 2007;17:370–82.
- Dreinhöfer K, Stucki G, Ewert T, Huber E, Ebenbichler G, Gutenbrunner C, et al. ICF core sets for osteoarthritis. J Rehabil Med. 2004;36(Suppl):75–80.
- Stucki G, Cieza A, Geyh S, Battistella L, Lloyd J, Symmons D, et al. ICF core sets for rheumatoid arthritis. J Rehabil Med. 2004;36(Suppl):87–93.